Smoother: A Unified and Modular Framework for Incorporating Structural Dependency in Spatial Omics Data

Jiayu Su[*1,2,3], Jean-Baptiste Reynier[1,4], Xi Fu[1,4], Guojie Zhong[2], Jiahao Jiang[5], Rydberg Supo Escalante[2], Yiping Wang[1,6], Luis Aparicio[1,2], Benjamin Izar[1,6], David A Knowles[2,3,7], Raul Rabadan[*1,2,4]

[1]*Program for Mathematical Genomics, Columbia University, New York, NY, USA*

[2]*Department of Systems Biology, Columbia University, New York, NY, USA*

[3]*New York Genome Center, New York, NY, USA*

[4]*Department of Biomedical Informatics, Columbia University, New York, NY, USA*

[5]*Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK*

[6]*Department of Medicine, Division of Hematology/Oncology, and Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY, USA*

[7]*Department of Computer Science, Columbia University, New York, NY, USA*

[*] Correspondence: js5756@cumc.columbia.edu, rr2579@cumc.columbia.edu

Abstract

Spatial omics technologies can help identify spatially organized biological processes, but existing computational approaches often overlook structural dependencies in the data. Here, we introduce Smoother, a unified framework that integrates positional information into non-spatial models via modular priors and losses. In simulated and real datasets, Smoother enables accurate data imputation, cell-type deconvolution, and dimensionality reduction with remarkable efficiency. In colorectal cancer, Smoother-guided deconvolution revealed plasma cell and fibroblast subtype localizations linked to tumor microenvironment restructuring. Additionally, joint modeling of spatial and single-cell human prostate data with Smoother allowed for spatial mapping of reference populations with significantly reduced ambiguity.

Keywords

Spatial omics, Spatial prior, Data imputation, Cell-type deconvolution, Dimensionality reduction, Reference mapping, Joint analysis of single-cell and spatial data

Background

From subcellular arrangement to tissue compartmentalization, the spatial structure in an organism is highly organized at all scales. This harmonious architecture regulates a diverse variety of biological processes, including embryonic development, neuronal plasticity, and the tumor microenvironment. Recent advances in spatially resolved omics technologies provide unique opportunities to study localization patterns of gene and epigenetic activities, as well as the dynamics of biological systems at cellular and tissue levels(1-5). While existing non-spatial omics analysis methods can be applied to spatial data, the neglect of positional information makes them inadequate in overcoming structured technical noise(6), let alone inferring

biologically meaningful spatial organization. Meanwhile, *ad-hoc* spatially aware models often hardcode neighborhood structures into task-specific algorithms(7-15), hindering adaptation to new applications without substantial modification. Even within the same application, models pretrained on one sample usually cannot be applied to another, as the neighborhood graph varies from sample to sample. Additionally, these models are also incompatible with non-spatial data, including the rich single-cell omics atlas datasets, and therefore cannot transfer knowledge learned from non-spatial modalities to enhance spatial analysis.

Physically adjacent spots or cells generally exhibit more similarity compared to distant pairs, with the similarity decaying with distance at different rates (**Supplementary Fig. 1**). Such patterns are consistently observed across tissues, technologies, and modalities — even in single-cell resolution data**(16)** (**Supplementary Fig. 1e**) and in super-resolved tumor microenvironment sections(17, 18) (**Supplementary Fig. 1d and f**). Permutation experiments further confirmed that the long-range similarity structure is not an artifact of contamination or signal bleeding (**Supplementary Fig. 1i**). Despite its universality, spatial dependency in omics data has yet to be formally described in a generic framework independent of downstream applications. Instead, existing algorithms developed for individual tasks often regard spatial variation as a task-specific property, unnecessarily restricting their generalizability to new applications. For example, Markov random field-based models like Giotto(7), BayesSpace(8), CARD(9) and BayesTME(10), although utilizing the same Bayesian message passing mechanism, each introduce structural dependency with unique, integrative, and non-sharable implementations. This practice becomes especially troublesome when the prior belief needs modification, for instance, to encode boundary information or scale to larger datasets. Similarly, graph-based

neural networks such as SpaGCN(11) and STAGATE(12) also incorporate spatial structure as a hard constraint and integral part of the model. While interactions between neighbors can be learned adaptively from the data, these models are essentially black boxes, leaving users with minimal control over over-smoothing and signal dilution.

Here we present Smoother, a unified and modular framework for integrating spatial dependency across applications. By representing data as boundary-aware weighted graphs and Markov random fields(19, 20), Smoother explicitly characterizes the dependency structure, allowing information exchange between neighboring locations and facilitating robust and scalable inference of cellular and cell-type activities. Through the transformation between spatial prior and regularization loss, Smoother is highly modularized and ultra-efficient, enabling the seamless conversion of existing non-spatial single-cell-based models into spatially aware versions. We demonstrate the versatility of Smoother by implementing and testing its performance on tasks including cell-type deconvolution and dimensionality reduction. Using simulated and real omics data of different modalities, our benchmarks highlight the substantial advantages of explicitly modeling and incorporating spatial dependencies. Furthermore, Smoother's soft regularization approach also supports spatially aware joint embeddings of data with and without neighborhood structure, potentially bridging the gap between spatial and single-cell analyses.

Results

Overview of the Smoother framework.

Smoother differs from existing models in that it treats data-specific dependencies as shared and reusable priors across downstream tasks, encouraging local smoothness on any spatial variable of interest (**Fig. 1, Methods and Supplementary Notes**). Inspired by penalized likelihood methods(21), we decouple the prior belief on dependency from the likelihood of a non-spatial data-generating model. This flexibility allows the same prior to be used in different models, and the same model to accommodate data with varying or even zero spatial structures. In addition, Smoother encodes boundary information that is often neglected in existing Bayesian methods (**Fig. 2**). Specifically, it first builds a spatial graph in which edges connecting physically adjacent locations are scaled and pruned using histological and transcriptomic similarities to remove undesired interactions (**Fig. 2a**). Through graph weighting, users may incorporate additional knowledge from other modalities, even though the actual region boundaries in omics data probably being less distinct (**Fig. 2e and Supplementary Fig. 2**), The spatial graph is then converted into a multivariate normal (MVN) prior with varying degrees of dependencies, along with an equivalent spatial loss that can be appended to non-spatial models (**Supplementary Fig. 3.** See **Methods** for detailed recommendations on constructing the prior).

As a simple showcase of Smoother, we considered feature smoothing and resolution enhancing, where variables at unobserved locations are inferred from a hidden Markov random field with the MVN prior(22) (**Fig. 1 left and Methods**). Unlike other data imputation algorithms(12, 15, 23), our model does not assume feature-level dependencies, making it applicable to a single feature of interest, especially non-expression data. We investigated whether it was possible to overcome data sparsity using spatial context alone. Gene signature scoring is a common approach to evaluate high-level activities of functionally associated gene sets(24), where

borrowing information across similar genes may introduce biases and artificially amplify the signal. By penalizing local variations over space, Smoother successfully mitigated dropout effects and improved the separation of localization patterns of cortical layer signatures in a human dorsolateral prefrontal cortex (DLPFC) dataset(25) (**Supplementary Fig. 4**). Furthermore, Smoother offers ultra-fast functionality, without requiring additional inputs like histological images(15), to enhance spatial resolution to any scale in seconds (**Supplementary Fig. 5**). While accuracy relies on the assumption that spatially adjacent spots share more similarity, this can be practically helpful when working with shallow sequencing depths. On a Slide-seqV2 dataset of human melanoma brain metastasis(17), we noted improved correlations in the activities of functionally connected modules, such as chemokine and interferon response, after smoothing and enhancing (**Supplementary Fig. 6**).

Smoother enhances cell-type deconvolution performance in simulated and real spatial omics data.

A common challenge with barcode-based spatial omics technologies that capture a mixture of cells at each location is to disentangle cell-type composition across space. Despite the many deconvolution methods(26-31) developed for spatial transcriptomics (ST) data, few recognize spatial dependency explicitly. Even for these methods, for instance CARD(9) and BayesTME(10), the spatial covariance is hard-coded and not transferable to other models. To fill this gap, we extended four non-spatial deconvolution models using Smoother to resolve the distribution pattern of cell types in a spatially informed manner (**Fig. 1 bottom**). These include non-negative least squares (NNLS), support vector regression (SVR)(32), dampened least squares (DWLS)(33), and log-normal regression (LNR)(34). To benchmark the effect of

Smoother, we first simulated ST data with distinct patterns, diverse degrees of spatial heterogeneity, and composition-independent structural noise corresponding to spot bleeding (**Supplementary Fig. 7, 14 and Methods**). In almost every simulation scenario, the inclusion of spatial context consistently yielded more accurate and realistic deconvolution results (**Fig. 3 and Supplementary Figs. 8-14**). Compartment boundaries became notably cleaner as real signals stood out against the prior while noise was smoothed out (**Fig. 3c and g**). This benefit proved highly robust to the selection of models, hyperparameters, and marker genes, while the magnitude shrinks as cell-type-independent spatial variation like spot bleeding grew (**Supplementary Figs. 15-17**). Surprisingly, we observed that CARD failed to take advantage of its own neighborhood modeling, and thus in some scenarios performed worse than a simple non-spatial model. Our results indicate that the parameterization of spatial structure used in CARD is suboptimal, underscoring the importance of a unified and modular framework for dependency representation.

We next applied all spatially aware deconvolution methods to analyze real spatial omics data of normal, developmental and cancer tissues. Descriptions of each dataset and preprocessing details can be found in the Methods section (**Methods**). We first evaluated deconvolution performance in detecting immune infiltration in a 10x Visium invasive ductal carcinoma dataset(8), where CD3 staining provided ground truth for the presence of T cells. Smoother-guided NNLS faithfully reflected overall T-cell distributions and accurately unveiled the invasive lymphocytic pattern near tumor borders, whereas CARD failed to detect any T-cell signal within the tumor (**Fig. 4a and Supplementary Fig. 18**). The advantage of spatial context modeling was further validated by the elevated correlation of CD3 staining with estimated abundance, outcompeting

baseline correlations between gene expression and protein staining (**Fig. 4b and c**). Secondly, to evaluate Smoother's ability to filter noise while maintaining meaningful boundaries, we performed deconvolution on a 10x Visium dataset of mouse brain(26) and estimated the abundances of 52 neural subtypes across brain locations simultaneously (**Fig. 4d**). As revealed by unsupervised clustering, Smoother notably reduced local discontinuities in cell-type composition, while preserving distinct tissue region boundaries (**Fig. 4e**). In particular, Smoother-based models sharply distinguished excitatory neuron subtypes in cortex and hippocampus, attenuating fluctuations within layers and subregions and producing more precise transitions compared to CARD (**Fig. 4f and g**).

Using a spatial-CUT&Tag dataset of mouse E11 embryo, we further assessed the generalizability of Smoother-based deconvolution across modalities (**Fig. 4h and Methods**). Spatial-CUT&Tag(4) profiles chromatin modification using antibodies against histone proteins including H3K27me3, H3K4me3, and H3K27ac. To align the epigenomic data with transcriptomic cell types, we performed deconvolution on epigenomics-based gene activity scores using scRNA-seq reference of 37 embryonic cell types(35) (**Supplementary Fig. 19**). These activity scores differ significantly from expression counts in scale, variability, and biological meaning. Moreover, some transcriptomic cell types may be epigenetically indistinguishable, thereby necessitating more robust deconvolution. Regularization has long been recognized as a solution to the multicollinearity problem(36). Consistently, Smoother salvaged the poor performance of the non-spatial model, producing biologically coherent embryonic compartmentalization (**Fig. 4i**). Close examination of the predicted cell-type composition showed that the Smoother-guided approach excelled in restoring the spatial organization of the embryo with minimal unsolved

background (**Fig. 4j**). In particular, only the spatially regularized NNLS model accurately mapped the cardiac muscle lineage to the heart region (**Supplemental Fig. 20-22**). The predicted spatial structures of glial (*Sox9*), neural tube (*Ascl1*), cardiac muscle (*Hand2*), and definitive erythroid (*Lmo2*) lineages were further validated by marker gene activity (**Fig. 4k**).

Smoother-guided deconvolution unveils distinct localizations of plasma cells and fibroblasts in mismatch repair-proficient colorectal cancer.

To further demonstrate the utility of Smoother-guided deconvolution in large-scale cancer ST datasets, we applied the method to a new Stereo-seq mismatch repair-proficient (MMRp) colorectal cancer (CRC) dataset with paired patient-derived scRNA-seq data(18). Utilizing gene modules from prior research(37), we identified eight major cell types and 16 functionally distinct subsets in the scRNA-seq reference, including three plasma cell subpopulations: namely IgG+, IgA+, and IgA+ FOS/JUN+ (**Fig. 5a and b**). *FOS* and *JUN* are B-cell receptor pathway modules and have been shown to be upregulated in tumor-infiltrating B cells in other solid tumors(38).

We then mapped these single-cell populations to spatial locations through deconvolution. Overall, the Smoother-guided model outperformed CARD in recapitulating the pathological structure, particularly at region boundaries (**Fig. 5c and d**). This facilitated the discovery that the three plasma populations resided in distinct regions: specifically, the IgG+ population were predominantly in the tumor region (lesion) whereas the IgA+ cells were in mucosa (**Fig. 5c and e**). We confirmed the differential localization of IgG+ and IgA+ plasma cells using marker gene expression (**Fig. 5f and Supplementary Fig. 23**). In addition, the lesion IgG+ and mucosa IgA+ spots exhibited divergent B-cell receptor V/C gene usage, with *IGLC2*, *IGLC3*, *IGLV3-1* and

IGKV4-1 all enriched in the tumor, suggesting a difference in the adaptive response between the two plasma cell types (**Fig. 5f**). Gene Ontology pathway analysis further emphasized this difference, with "B-cell activation" enriched in the IgG+ plasma cell sections and "Defense response to bacterium" in the IgA+ plasma cell sections (**Fig. 5g**). Our observation aligns with the established role of IgA+ plasma cells in colorectal mucosal tissues(39), as well as numerous reports of antibody class switching to IgG in the CRC tumor microenvironment(40), which has high potential for diagnostics(41, 42) and therapeutics(37, 43). In stark contrast, CARD incorrectly predicted all plasma cell clones to be IgG+, including in the mucosa region (**Supplementary Fig. 23**).

Additionally, Smoother-guided deconvolution revealed three phenotypically and spatially distinct fibroblast populations, characterized by previously reported ADAMDEC1+/CCL8+, CXCL14+, and matrix transcriptional programs(37), all validated by marker gene expression (**Supplementary Fig. 24**). ADAMDEC1+/CCL8+ and CXCL14+ fibroblasts were preferentially co-localized with the IgG+ plasma cells (**Fig. 5h**), in accordance with the observed presence of CXCL14+ cancer-associated fibroblast and ADAMDEC1+/CCL8+ fibroblast in MMRp CRC(37). Recent work has demonstrated *ADAMDEC1*-driven fibroblastic matrix remodeling in response to inflammation(44), which explains partially the enrichment of collagen and extracellular matrix-related pathways in IgG+ plasma cell sections (**Fig. 5g**). Conversely, CARD again failed to delineate these three fibroblast populations (**Supplementary Fig. 23 and 24**). We repeated the deconvolution analysis on a paired distant normal tissue but observed no differential localizations of plasma cell and fibroblast subpopulations (**Supplementary Fig. 25**).

Smoother-guided dimensionality reduction and integration of spatial and single-cell data. Learning an informative low-dimensional representation is crucial for understanding the biological dynamics underlying noisy omics data. Smoother's ability to impose structural dependencies via a versatile loss function allows us to generalize existing non-spatial dimensionality reduction methods to spatial omics data (**Fig. 6a, Methods and Supplementary Notes**). This is further strengthened by the contrastive extension of the spatial loss that better separates distant locations and averts the collapse of embeddings (**Methods**). As a proof of concept, we first developed a spatially regularized principal component analysis (PCA) model(45) and applied it to the human cortex DLPFC data(25). Across all 12 samples, the inclusion of spatial loss consistently improved performance when reconstruction and regularization were balanced, with contrastive loss further strengthening its robustness (**Supplementary Fig. 26a and b**). These enhancements are manifest in the 2D visualization where outer cortical layers are increasingly separated from inner layers and the white matter with stronger regularization (**Supplementary Fig. 26c**). Nevertheless, existing graph-based autoencoders like STAGATE(12) and SpaceFlow(13) achieved even better separation, indicating the intrinsic limitation of PCA as a linear encoder. To dissect the benefit of spatial modeling from simply having a broader parameter space, we replaced the graph attention module in STAGATE with fully connected layers and again incorporated spatial loss during training (**Supplementary Fig. 26d-h**). By increasing the strength of spatial regularization, we recovered 50% - 90% of the performance as measured by embedding consistency (Silhouette score) and clustering accuracy (adjusted rand index) (**Supplementary Fig. 26d-g**). Under 2D UMAP visualization, the space progressively unfolded, and transitions between layers evolved from being non-existent to becoming clear

(**Supplementary Fig. 26h**). Together, our results support the potential of spatial loss as a fast and versatile alternative strategy to instill spatial awareness irrespective of model architecture.

This modular approach, especially the separation of neighborhood constraints from the autoencoder, offers unique advantages. Through Smoother, we can now adapt pretrained single-cell based non-spatial models to spatial data by fine-tuning over a new spatially aware objective, providing an approach to transfer knowledge from single-cell atlases to spatial omics data (**Fig. 6a**). Specifically, we designed SpatialVAE from SCVI, a prominent deep generative variational autoencoder (VAE) for scRNA-seq data analysis(46, 47). When integrating scRNA-seq datasets, the conventional SCVI workflow includes training a reference model on large atlas datasets, fine-tuning on the query data, and generating joint embeddings of both for label transfer. Here we focused on annotating a human prostate Slide-seqV2 data(48) using a single-cell reference from the Tabular Sapiens(49). After downloading the pretrained model (**Fig. 6b**) from the SCVI model hub, we fine-tuned the model on the spatial data while ignoring coordinates (RNA-only) to remove batch effects. Although SCVI ranks among the top performing data integration tools(50), the Slide-seqV2 data remained distinctly separated from the rest of single-cell data in the joint embedding space (**Fig. 6d**), likely due to its low sequencing depth (**Fig. 6c**). Consequently, most spots were marked as ambiguous, that is, not close to any single-cell clusters within the uncertainty threshold (**Fig. 6e**). In contrast, SpatialVAE's spatially aware refinements significantly reduced batch effects (**Fig. 6f**), enabling cell label transfer with dramatically reduced ambiguity (**Fig. 6g**). Detailed examination revealed a balanced tradeoff between reconstruction precision and spatial coherence of the embeddings, suggesting that the spatial loss potentially acts via directing the model to focus on spatially consistent features over batch-

specific technical noise (**Supplementary Fig.27a**). Using the prostate joint embeddings, we also corrected an annotation mistake in the original publication where cell type labels were mapped to the Slide-seqV2 data using RCTD deconvolution(28). While initially labeled as fibroblasts, the stromal population were indeed primarily composed of smooth muscle cells (referred to as pericytes in(48)) according to SpatialVAE's prediction (**Supplementary Fig. 27b**). This was confirmed by marker gene expression (**Supplementary Fig.27c-d**).

Discussion

The inherent neighborhood dependency of spatial omics data motivated us to develop an efficient approach to introduce spatial structure across single-cell applications. Grounded in Bayesian inference and penalized likelihood methods, Smoother imposes regularization at minimal computational burden, particularly by leveraging the sparse nature of the neighborhood graph. In practice, tasks such as enhancing resolution and deconvolving tens of thousands of spots can typically be completed within seconds on a standard personal laptop. The efficiency positions Smoother as an apt solution for the growing demands to explore spatial omics techniques with larger field-of-view, higher resolution, and increased throughput. One caveat to Smoother is that the optimal strength of spatial dependencies, which reflects prior assumptions on the data, is usually agnostic beforehand in downstream applications. Still, we have demonstrated that Smoother is remarkably robust with respect to hyperparameters. It may also be possible to fine-tune the smoothing strength in specific tasks using cross validation and empirical Bayes approaches(51). Collectively, Smoother offers a scalable, and versatile solution to enhance a wide range of tasks including data imputation, deconvolution, and dimensionality reduction. Under this framework, the spatial loss can be melded with any optimization-based non-spatial

model, pretrained or otherwise, endowing it with spatially awareness. In light of the ubiquity of structural dependencies, we envision that Smoother may be readily extended to even more applications, such as trajectory inference and cell-cell communication, paving the way for new biomedical discoveries in developmental and disease settings.

Conclusion

In this study, we introduce Smoother, a powerful and adaptable computational approach designed to integrate spatial structure into omics data analysis. Through spatial priors and losses, Smoother provides a streamlined and efficient way to rewire existing single-cell-based models for spatially informed applications, including feature smoothing, resolution enhancing, cell-type deconvolution, and dimensionality reduction. Benefiting from its robustness, Smoother-regularized deconvolution accurately mapped transcriptomic cell types to spatial epigenomics data. When applied to colorectal cancer sections, it further revealed tumor-associated localizations of plasma cell and fibroblast subpopulations. Furthermore, Smoother's compatibility with non-spatial data allowed for the spatially informed integration of human prostate data, facilitating cell type prediction with markedly lower ambiguity.

Methods

Overview of Smoother

Smoother is a two-step framework that extracts prior dependency structures from positional information and integrates them into non-spatial models for downstream tasks to encourage local smoothness. We use $Y_{G\times S}$ to denote the spatial omics data of G genes (or the corresponding features in other modalities) at S locations, and $L_{F\times S}$ the spatial metadata of F spatial features (e.g., image feature extracted from histology) at S locations. The first step of Smoother is to construct a spatial graph G = (V, E, W) where the node set V represents locations and both the edge set E and edge weights W are functions of $L_{F\times S}$. Then, Smoother computes a covariance structure $\Sigma_{S\times S}$ from the graph G and imposes it through the spatial prior on $Y_{G\times S}$ or other variables of interest. For a comprehensive explanation of Smoother, see below and the **Supplementary Notes.**

Construction of spatial priors and losses.

Smoother draws from the concept of spatial stochastic process and an extensive body of spatial priors in image processing and geographic data analysis(22, 52, 53). Intuitively, a spatial stochastic process lets us model global dependency characteristics by assuming stationarity and specifying the local interactions of spatial random variables. Here we represent spatial connectivity using a weighted undirected graph G = (V, E, W) where physically adjacent locations V are connected by edges E with varying strengths W. The adjacency matrix $W_{S\times S}$, which is also called the spatial weights matrix of the underlying spatial process, maps out the connectivity between spots based on physical distance. Here $w_{i,j} = 1$ if spot i and j are mutual k-nearest neighbors, otherwise 0. For hexagonal grids like the 10x Visium chip, k is set to 6. To

encode domain boundaries, we further scale $W_{S\times S}$ using histological or transcriptomic pairwise similarities (soft-scaling), or manual domain annotations (hard-scaling):

$$W'_{S\times S} = W_{S\times S} \odot D_{S\times S}^{soft,hard}$$

where $D_{i,j}^{soft}$ is the similarity between spot i and j, and $D_{i,j}^{hard}$ is the binary indicator of whether spot i and j belong to the same domain. In practical applications, hard-scaling domains can be defined from histology image segmentation(54), transcriptomic clustering, or expert pathological annotations. For soft scaling, we extract per-spot gene expression or histological features and compute $d_{i,j}$ as the pairwise similarity in a PCA-reduced space. If scaling by transcriptomics, the first 10 PCs of gene expression and cosine similarity (which is approximately the cosine similarity of Z-scores of full gene expression) are used by default and negative similarities are clamped to zero. If scaling by histology, the feature vector of a spot is the first three PCs of the concatenated RGB values of pixels in the square circumscribed about the spot, and similarity is converted from the Euclidean distance by a Gaussian kernel with band width 0.1 in the normalized PC score space. Specific choices on similarity metrics usually do not have a strong impact on the resulting prior. Empirically, we found the scaling of gene expression to be helpful in maximizing dissimilarity between disparate neighbors.

Subsequently, Smoother translates the spatial weights matrix $W_{S\times S}$ into a covariance structure $\Sigma_{S\times S}$ according to assumptions on the underlying stochastic process (see **Supplementary Notes**). The covariance is then introduced to any spatial variable of interest, X_S, through a multivariate normal (MVN) prior. For example, in a conditional autoregressive (CAR) process where the graph G describes a Gaussian Markov random field of X_S:

$$x_i \mid x_{\setminus i} \sim N\left(\rho\Sigma_{j\neq i}\, w_{ij}\, x_j, \sigma_i^2\right),$$

the joint distribution of X_S is a zero-centered MVN distribution with covariance $\Sigma_{S\times S}$, a smoothing prior that can be imposed on X_S in downstream tasks:

$$\Sigma_{S\times S} = diag(\sigma_i^{-2})(I - \rho W)^{-1}$$

$$P(X_S) \propto exp\left(-\frac{1}{2} X_S^T \Sigma_{S\times S}^{-1} X_S\right).$$

Here ρ is the autocorrelation parameter to make $\Sigma_{S\times S}$ positive semi-definite and to control the decay rate of covariance over distance (**Supplementary Fig. 3**), which can be selected by examining the decay pattern of pairwise similarity (**Supplementary Fig. 1**). Since $\Sigma_{S\times S}$ is constructed from a boundary-aware graph, the above MVN prior provides a unique channel for neighboring locations to share information while still preserving boundaries.

Smoother offers five different yet related spatial processes: CAR, SAR (simultaneous autoregressive), ICAR, ISAR, and SMA (spatial moving average). Specifically, CAR and SAR are equivalent upon transformation, and ICAR and ISAR are the weights-scaled versions so that the autocorrelation parameter ρ falls in [0, 1]. By adjusting ρ, these models can achieve parallel regularization effects. Based on numerical considerations, we typically recommend using ICAR with varying ρs (or ISAR with smaller ρs) to accommodate data with diverse neighborhood structures. For instance, 'ICAR (ρ =0.99)' for data with clear anatomy and 'ICAR (ρ =0.9)' for tumor data. SMA is generally not recommended since the resulting inverse covariance matrix tends to be less sparse, potentially slowing down computation.

To render Smoother compatible with existing non-spatial methods, we propose a quadratic regularization loss proportional to the density function of a zero-centered MVN with covariance $\Sigma_{S\times S}$:

$$L_{sp}(X_S; \Sigma_{S\times S}) = X_S^T \Sigma_{S\times S}^{-1} X_S.$$

Essentially, for any given model with a loss function L_m, it is possible to morph the model into a spatially aware version by minimizing a new joint loss function:

$$L_{joint}(X_S) = L_m(X_s) + \lambda_{sp} L_{sp}(X_S; \Sigma_{S\times S}).$$

The spatial loss term L_{sp} regularizes local fluctuations in X_s and makes the inference robust to technical noise. It can be shown that optimizing the new objective is equivalent to finding the maximum a posteriori (MAP) estimator under the spatial prior, and λ_{sp} can be viewed as the strength (inverse variance) of the prior. Most importantly, since L_{sp} is separated from the model loss, the same model can jointly accommodate data with or without neighborhood structures.

In addition, we implement a contrastive extension of the spatial loss to increase the penalty for pulling distant spots too close, ensuring that the inference does not collapse into trivial solutions. This is done by shuffling spot locations and producing corrupted covariance structures as negative samples,

$$L_{csp} = X_S^T \Sigma_0^{-1} X_S - \frac{1}{T}\sum_t X_S^T \Sigma_t^{-1} X_S$$

where Σ_0 is the covariance derived from the correct spatial graph and $\Sigma_t, t \in [1, T]$ are those from corrupted graphs.

Data imputation and resolution enhancement

Using local contexts in the prior dependency structure, Smoother is capable of smoothing and imputing spatial omics data at unseen locations for any single spatial random variable of interest

X_S. For simplicity, we assume the variable follows a hidden Markov random field model with technical noise being Gaussian iid. This implies that the observation Y_S follows the model:

$$y_s \sim N(x_s, \sigma_0^2)$$

$$X_S \sim MVN(0, \sigma_1^2 \Sigma_{S \times S})$$

where σ_0^2 and σ_1^2 are the variance of observation and prior, respectively. Similarly, we can reparametrize the above problem and find the MAP estimator of X_S by solving the following optimization task with a given spatial loss:

$$\tilde{X}_{MAP} = argmin_X \|Y_S - X_S\|^2 + \lambda_{sp} X_S^T \Sigma_{S \times S}^{-1} X_S$$

where $\lambda_{sp} \propto \sigma_0^2 / \sigma_1^2$ determines the strength of regularization. This is a special case of Tikhonov regularization and is akin to the weighted average filter commonly used for image smoothing. Note that the first L2 term, corresponding to the reconstruction error, can be replaced by other likelihood-based or more sophisticated losses in deep generative models for non-Gaussian variables. When part of the data is missing, the objective function is similar except the first reconstruction term is computed only at observed positions. As $\Sigma_{S \times S}$ is predefined independently of the observation, Smoother can impute the latent value at arbitrary locations and thus increase the spatial resolution.

Cell-type deconvolution

We define spatial deconvolution as the problem of inferring cell-type abundances at each location from the observed omics data, with or without cell-type reference information from external data. The task is especially relevant for spatial techniques with limited resolution, including 10x Visium, spatial-CUT&Tag(4) and spatial-ATAC-seq(5), where each profiled location might consists of cells from multiple cell types. The deconvolution model is usually

determined by the generative model of the observations $Y_{G\times S}$. Most deconvolution methods assume a linear relationship between observation and cell-type abundance:

$$Y_{G\times S} \sim N(g(R_{G\times C}X_{C\times S}), \sigma_0^2).$$

Here, $Y_{G\times S}$ is the observed activities of G features at S spots, $R_{G\times C}$ is the expected reference activities of G features in C cell types, $X_{C\times S}$ is the abundance of C cell types at S spots, σ_0^2 is the sampling variability, and g is the data generative function that introduces additional noise such as location-specific biases. Without loss of generality, we follow the same linearity assumption and extend existing deconvolution models to leverage neighborhood information by imposing the spatial prior on the abundance of each cell type:

$$X_{c:} \sim MVN(0, \sigma_1^2 \Sigma_c).$$

Here, Σ_c is the prior covariance structure, and $1/\sigma_1^2$ represents the strength of prior. When the reference $R_{G\times C}$ is known, usually from a paired single-cell dataset, we solve $X_{C\times S}$ by minimizing the following regularized factorization problem:

$$\tilde{X}_{MAP} = argmin_X \, L_{deconv}(Y, RX) + \lambda_{sp} \sum_c X_{c:} \Sigma^{-1} X_{c:}^T$$

where $L_{deconv}(Y, X\beta)$ is the loss specified by the corresponding deconvolution model. We implemented four spatially aware deconvolution models including non-negative least squares (NNLS), support vector regression (SVR), dampened least squares (DWLS), and log-normal regression (LNR). Further details can be found in the **Supplementary Notes**. When the reference is unknown, the above deconvolution can be solved via matrix factorization, which is also a special case of the dimensionality reduction task, as will be discussed in the next section.

Dimensionality reduction

The goal of dimensionality reduction is to infer a condensed low-dimensional representation retaining the data's essential characteristics. For spatial omics data, latent dimensions should ideally represent continuous dynamics in space. We frame the dimensionality reduction task using a general autoencoder model,

$$Y_s \xrightarrow{E_\theta} Z_s \xrightarrow{D_\phi} Y_s{}'$$

where E_θ is the decoder that projects omics data $Y_s \in R^G$ at the location s with G features onto a low-dimensional space R^H with H hidden dimensions, and D_ϕ is the decoder that projects the hidden embedding $Z_s \in R^H$ back to the original space.

Using Smoother, we can again impose prior on the hidden embedding to get more coherent representation. For any auto-encoder model with parameter θ, ϕ and reconstruction loss function $L_m(Y_s, \theta, \phi)$, Smoother regularizes the hidden representation $Z_s(Y_s; \theta) = E_\theta(Y_s)$ using a spatial loss L_{sp} and solves a new joint objective

$$\hat{\theta}, \hat{\phi} = argmin_{\theta,\phi}\ L_m(Y, \theta, \phi) + \lambda_{sp} L_{sp}(E_\theta(Y); \Sigma_{S\times S}) + \lambda_c L_c(\theta, \phi)$$

where $L_c(\theta, \phi)$ is the additional soft constraint loss on model parameters, if any. For linear dimensionality reduction models, including non-negative matrix factorization (NMF)(55), principal component analysis (PCA)(45), and independent component analysis (ICA)(56), the encoder E_θ and decoder D_ϕ are matrix multiplication operators and can be resolved via matrix factorization:

$$Y_{G\times S}{}' = W^{\phi}_{G\times H}\left(W^{\theta}_{H\times G} Y_{G\times S}\right)$$

$$\widetilde{W^\theta}, \widetilde{W^\phi} = argmin_{W^\theta, W^\phi}\ L_m\left(Y, W^\phi\left(W^\theta Y\right)\right) + \lambda_{sp}\ L_{sp}\left(W^\theta Y\right) + \lambda_c L_c\left(W^\theta, W^\phi\right).$$

The apparent analogy between the above linear autoencoder model and linear deconvolution suggests that cell-type abundance itself can be viewed as the hidden factor. When the deconvolution reference $R_{G\times C}$ (i.e., $W_{G\times H}^{\phi}$ in the autoencoder) is unknown, factorizing $Y_{G\times S}$ is both a dimensionality reduction task and a reference-free deconvolution task. Further semi-supervised techniques can be applied to maximize the distinguishability of inferred latent states (cell types) using known marker genes.

In this study, we developed corresponding spatial versions of PCA, vanilla deep autoencoder, and Variational Autoencoder (VAE) within the Smoother framework. We assume all hidden dimensions to be independent and regularize them simultaneously. Briefly, the PCA model has symmetric encoder and decoder, requires latent dimensions to be orthogonal and uses L2 norm to measure reconstruction error. The deep autoencoder model removes the symmetry and orthogonality constraints, introduces nonlinearity, and incorporates an orthogonal loss to impose soft constraints on the latent embedding. In a VAE model that describes the data generative process and learns the distribution of latent representations, regularization is by default applied to the mean of the inferred latent distribution.

Model implementation

Smoother is publicly available as a Python package (https://github.com/JiayuSuPKU/Smoother/). Models involved in this study are implemented using PyTorch(57) and all optimizations are solved via PyTorch's gradient-based optimizers (by default Adam for deconvolution and SGD for dimensionality reduction). For convex problems, an alternative Smoother implementation via

CVXPY(58) is also available. For VAE models, we adopted the VAE implementation from the Python package scvi-tools(47).

Preprocess spatial omics data

Unless otherwise noted, we downloaded spatial omics datasets from the SODB database(59) where the data were preprocessed following the Scanpy workflow(60). For data not available through SODB, we used the default Scanpy preprocessing workflow. In **Supplementary Fig. 1**, we used the first 20 PCs of log-normalized expression data to calculate pairwise similarity decay. Preprocessing details of the spatial-CUT&Tag data are provided below in the corresponding subsection. Spatial priors across applications were constructed by default using ICAR with ρ = 0.99 and transcriptomic soft-scaling, unless otherwise specified.

Recover spatial patterns using Smoother-guided imputation and resolution enhancement

Human dorsolateral prefrontal cortex (DLPFC) dataset

For each cortical layer, we calculated gene signature scores using the 'scanpy.tl.score_genes' function based on the expression of the top 20 marker genes ranked by log fold change. The spatial prior was constructed under the ICAR model with ρ = 0.99 and $\lambda_{sp} = 1$. For imputation, we randomly masked out certain proportions of spots in a slide and allowed the target variable to vary in both observed and unobserved locations. For resolution enhancement, we added new spots according to the desired new resolution through midpoint interpolation and ran imputation on the new slide. Variable values at observed locations were fixed during enhancing.

Human metastatic melanoma Slide-seqV2 dataset

We preprocessed the Slide-seqV2 data of melanoma brain metastasis (MBM) and extracranial melanoma metastasis (ECM) and the paired scRNA-seq data following the original publication(17) using the R package Seurat(61). Functional signature scores were calculated based on the expression of genes involved in the corresponding pathway using Seurat's 'AddModuleScore' function. We built an ICAR prior with $\rho = 0.9$ and $\lambda_{sp} = 1$ to impute function scores individually and to enhance the spatial resolution. Pairwise Pearson correlation between functional scores was calculated to evaluate the benefits brought by data imputation.

Evaluate cell-type deconvolution performance using simulation

Simulate ST datasets with spatial patterns

To generate synthetic ST datasets for benchmark, we followed a modified procedure adapted from the cell2location paper(26). These modifications allowed us to: (1) assign arbitrary spatial patterns to zones (co-localized cell-type groups) and (2) introduce additional noise for cell-type-independent dependencies (e.g. spot bleeding). In brief, we initially overlaid designated patterns with a two-dimensional Gaussian process to generate the per-zone abundance values in space, then assigned cell types to these patterns. Next, we sampled gene expression profiles at each location from a scRNA-seq reference according to the simulated cell-type composition. The ST data is further combined with various sources of noise, including lateral diffusion where each spot shares a certain proportion of mRNA to its neighbors directly, and multiplicative per-gene gamma noise (by default shape = 0.4586, scale = 1/0.6992) for sampling error. The simulation code is provided as a stand-alone tool within the Smoother package to facilitate future benchmarking on related tasks.

Model comparison and evaluation

In the study, we focused our benchmarking on CARD, the only published spatially aware deconvolution method shown to be superior to other existing methods(9). However, Smoother is compatible with different deconvolution strategies. Any method, including the DWLS(33), nu-SVR(32), and LNR(34) re-implemented in this study, may be seamlessly transformed to take advantage of neighborhood information. CARD is a non-negative linear factorization model that introduces spatial dependencies through a conditional autoregressive prior on cell-type abundances, which is mathematically similar to the Smoother-guided NNLS model. We adhered to the tutorial and ran CARD with default parameters. In certain simulation scenarios, a truncated reference signature matrix with fewer genes was used as input, bypassing CARD's internal reference construction process. This was to reflect potential discrepancies between the external reference and observed ST data, and to separate the impact of reference quality from algorithm performance. For the non-spatial CARD model, we set ρ to zero, effectively disabling any spatial interactions.

The performance of deconvolution depends on the distribution properties of the input data. For example, RNA-seq counts are typically skewed and, if left uncorrected, can bias the estimation against lowly expressed genes and rare cell types(34). One practical solution is to perform deconvolution on the logarithmic scale, i.e., replacing Y and X with $log1p(Y)$ and $log1p(X)$. Although this approach is not physically sound, it has been shown to significantly improve model performance, so does the square root scaling $\sqrt{Y}$ and $\sqrt{X}$ albeit to a less extent (data not shown). For benchmarking purposes, unless otherwise stated, we supplied log-scale ST data to NNLS, DWLS, and nu-SVR and raw-scale ST data to LNR and CARD. The reference

expression matrix was computed by averaging the normalized expression of marker genes across all cells of a given cell type from an external scRNA-seq dataset, followed by log-transformation where necessary. For more general cases, such as in epigenomics deconvolution where the reference and observation data may not be on the same scale or even from the same modality, users may include whatever preprocessing steps that best fit the dataset.

For Smoother-guided models, we constructed the spatial prior using ICAR (ρ = 0.99, which was also the optimal ρ in CARD) and scaled the graph using transcriptomic similarity. The strength of the spatial prior λ_{sp} is set to 0 for all non-spatial baseline models, 1 for spatially aware NNLS, SVR, LNR, and 3 for spatially aware DWLS to adjust for the inflated model loss after scaling. This is not necessarily the best performing setting as revealed in the parameter sensitivity analysis (**Supplementary Fig. 15-17**). Nevertheless, we fixed the strength across all benchmarks since the benefit of spatial context is rather robust. Cell-type abundances were set to be non-negative in all models and were normalized to output the final cell-type proportions at each spot.

Benchmark deconvolution performance using the simulated doughnut-shaped data

We obtained the single-nucleus RNA-seq data of mouse brain (5705STDY8058280, 5705STDY805828) along with cell type annotations for each cell from https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-11115/, and only kept the 15 most abundant neural subtypes for simulation. The dataset was further split into two, one to simulate ST data and the other as the deconvolution reference. We generated ST datasets on a 50x50 grid containing five high-density (average of four cells per occupied spot) and 10 low-density (0.4 per occupied spot) cell types. Each cell type was assigned to a unique but overlapping doughnut-

shaped distribution pattern. Cell-type-independent spatial dependencies were introduced by sharing 0/10%/50% of mRNA counts at each location to its first-degree neighbors (i.e. spot bleeding). For each simulated dataset, we performed deconvolution using four sets of genes: (1) The union of the top 20 marker genes for each cell type (n=283), (2) the union of the top 50 marker genes (n=656), (3) all discriminative informative markers genes whose log2(fold change) is larger than 1 in one and only one cell type (n=2693), or (4) informative genes selected by the CARD model (n=7972). This in total brings 12 scenarios, each is replicated 10 times. All preprocessing and differential expression analysis steps were performed using the Python package Scanpy(60).

We evaluate deconvolution performance using three metrics: (1) The mean square error and (2) the Pearson correlation between ground truth and the estimated proportion per cell type, and (3) the average precision score of binary prediction of whether a cell type is present at a given location (abundance >= 0.05) using the Python package Scikit-learn(62). Results across cell types and experiments (replicates) are aggregated together for boxplot and barplot visualizations.

Benchmark deconvolution performance using the sci-Space data of mouse embryo

We obtained the sci-Space data(63) of 14 mouse embryonic slides from GEO under the accession number GSE166692. Each single cell in the sci-Space data is labeled with an approximate spatial coordinate (~200 µm pitch). Cells of low-quality or from rare cell types (less than 100 cells in any slide) were filtered out. For the remaining eight cell types, we averaged the expression of 298 marker genes defined in the original paper to compute the deconvolution reference. To generate ST data of the same resolution, we pooled together all cells with the same

barcode and resampled cells in each spot to adjust for the uneven density of captured cells. The actual cell count at each location is determined by a gamma-Poisson distribution with an expectation of 10. Since most coordinates have only one associated cell, the resampling was executed with replacement from both the current spot and adjacent locations (with half probability), effectively increasing the diversity at each spot. To further challenge the deconvolution, we down-sampled the synthetic ST data to a maximum of 5,000 UMI per spot and added multiplicative per-gene gamma noise with the mean around 0.35.

Evaluate cell-type deconvolution performance using real spatial omics datasets

Benchmark deconvolution performance using the breast tumor dataset with staining

We obtained the invasive ductal carcinoma (IDC) data(8) with DAPI and anti-CD3 staining from 10x Genomics at https://support.10xgenomics.com/spatial-gene-expression/datasets and performed deconvolution to evaluate T-cell infiltration. The reference gene expression of nine cell types, including one tumor epithelial type, was computed using a scRNA-seq atlas of 26 breast cancer samples(64). For Smoother-guided models, we selected the top 20 unique differentially expressed genes for each cell type using Seurat(61) (180 genes in total) as the input for deconvolution. For CARD, we provided the scRNA-seq data with all genes as input and run with default parameters. Performance was first evaluated by measuring the Spearman correlation between the CD3 immunofluorescence (IF) intensity and the T-cell deconvolution proportion over all spots on the slide. We then investigated if these models could reliably reveal T-cell infiltration in tumor regions (labeled as invasive carcinoma regions in the original publication), a question of paramount importance for immunotherapy research. Specifically, we fitted a two-component Gaussian mixture model on the CD3 intensity and predicted 377 out of the 2078

spots in the tumor regions to be T-cell positive. Results were visualized as pie charts in **Supplementary Fig. 18**.

Visualize the smoothing effect on region boundaries using the 10x Visium mouse brain data

We obtained the 10x Visium dataset of mouse brain(26) and the paired snRNA-seq reference from ArrayExpress under accession IDs E-MTAB-11114 and E-MTAB-11115. The reference expression matrix for deconvolution was calculated for 52 neural subtypes after removing seven unknown or low-quality cell types. For Smoother-guided methods, the union set of the top 20 marker genes for each cell type discovered by Scanpy(60) was used as the input for deconvolution. For CARD, all QC-passed genes were supplied, and the reference matrix was estimated by CARD. To visualize brain regions with distinct cell-type compositions, we first built a KNN graph based on the the first 20 PCs of inferred cell-type proportions, then applied Leiden clustering using Scanpy(60) and aligned the resulting clusters across methods via linear sum assignment. Clustering resolution was manually adjusted for different deconvolution methods so that the number of clusters generated per method was approximately the same.

Spatial mapping of scRNA-seq defined cell types to the spatial-CUT&Tag data of mouse embryo

We obtained the spatial-CUT&Tag data and preprocessing scripts from the original publication(4). Specifically, the R package ArchR(65) and the 'getGeneScore_ArchR.R' script were used to compute gene activity scores for the top 500 variable genes using the same set of parameters. We removed spurious spots with outlier average gene activity (more than three standard deviations away from the mean), which lined up in one column or row on the grid and were considered as technical artifacts. The remaining spots were used in all downstream analyses.

For deconvolution, scRNA-seq data of E11 mouse embryo was obtained from Mouse Organogenesis Cell Atlas (MOCA)(35) and processed as described in the spatial-CUT&Tag paper. Specifically, we used the script 'integrative_data_analysis.R' to read and preprocess the subsampled (100k cells) data and Seurat's 'FindVariableFeatures' function to identify the top 500 variable genes. We then computed the pseudo-bulk expression of variable genes for cell types annotated in 'Main_cell_type' group and used it as the deconvolution reference. Spatial domains of cell types were determined and visualized using the same clustering strategy developed in the mouse brain analysis above.

Deconvolution analysis of the Stereo-seq data of mismatch repair-proficient colorectal cancer

Re-analyze and annotate the paired scRNA-seq data

We downloaded processed Stereo-seq and paired scRNA-seq data(18) from CNGB Nucleotide Sequence Archive under accession ID CNP0002432. The data was further processed as described below using Scanpy. After filtering out outliers ('n_genes_by_counts >= 2500' or 'pct_counts_mt' >= 5), we first clustered the single-cell reference into eight major cell types (T cell, B cell, mast cell, macrophage, plasma cell, epithelial cell, endothelial cell, and fibroblast) and confirmed their identities via marker genes. To identify functional subtypes, the three most abundant populations (T cells, plasma cells, and epithelial cells) as well as fibroblasts were separately re-clustered after regressing out the effect of sample donor. Each subtype was annotated according to the gene modules reported in (37). In particular, we re-discovered three plasma cell subsets (IgG+, IgA+, and IgA+ FOS/JUN+), three epithelial subsets (ADH1C, FOS/JUN, and TFF3), and three fibroblast subsets (CCL8, CXCL14, and Matrix), all with clear

markers. These subpopulations were either missed or misclassified in the original publication (18), likely due to batch effects.

Deconvolute Stereo-seq tumor and normal sections of patient P19

We extracted raw counts of the bin100 (50µm × 50µm) Stereo-seq data from the tumor and the paired distant normal tissue samples of patient P19 for separate deconvolution. For Smoother-guided deconvolution, the reference profile was the log averaged expression of the top 50 marker genes for the 16 subtypes (695 genes in total). For CARD, the reference was based on the same set of genes without log-scaling the expression. The default CARD-constructed reference contained much more genes (15025 genes in total) and yielded worse results (data not shown). The 'DWLS (+ spatial loss)' model shown in **Figure 5** was regularized using the ICAR prior with $\rho = 0.99$ and $\lambda_{sp} = 3$. Spatial domains of cell-type proportions were determined and visualized using the same clustering strategy (i.e. leiden clustering after PCA transformation) described in the mouse brain analysis above.

Differential localization of the IgG+ and IgA+ population in the tumor section

We divided spots into IgA- and IgG-specific groups based on the log ratio between 'PC_IgA' and 'PC_IgG' proportions estimated from the 'DWLS (+ spatial loss)' model (log ratio thresholds: >0.5 and <-0.5). For robustness, noisy proportions less than 0.01 were removed and a pseudo-count of 0.01 was added to all spots when calculating the log ratio. We then computed differentially expressed genes between the two groups (Wilcoxon test, p<=0.01) and further identified enriched 'GO:BP' pathways in each region using the function 'scanpy.queries.enrich'.

For colocalization analysis, we again removed noisy proportions (<0.01) and calculated statistical significance using the Wilcoxon test.

Joint embedding of Slide-seqV2 human prostate data and the Tabula Sapiens prostate scRNA-seq reference using SpatialVAE

We acquired raw Slide-seqV2 count data(48) of healthy human prostate samples and the associated annotations from https://github.com/shenglinmei/ProstateCancerAnalysis. The pretrained reference SCVI model and the training scRNA-seq data from the Tabula Sapiens prostate were downloaded from SCVI model hub at https://huggingface.co/scvi-tools/tabula-sapiens-prostate-scvi. Adhering to the conventional SCVI data integration workflow, we first fine-tuned the pretrained model on the Slide-seqV2 data with 'unfrozen=True' for 100 epochs to mitigate batch effects. This step also updated the latent representation of the single-cell reference. Model convergence was confirmed by inspecting the evidence lower bound (ELBO). For spatially informed fine-tuning, we converted the SCVI model into a new SpatialVAE model with spatial loss (ICAR, $\rho = 0.99$, $\lambda_{sp} = 0.01$). The strength λ_{sp} was selected to balance the spatial loss and the reconstruction accuracy (measured by log likelihood). A SpatialVAE model shares the same architecture and initial model parameters with the baseline model, which can be either the Tabula Sapiens reference model or the RNA-only fine-tuned model. In the study, we initialized SpatialVAE from the updated model and further fine-tuned with respect to the new objective for another 100 epochs. This was mainly to highlight the tradeoff between reconstruction accuracy and spatial consistency. Skipping the RNA-only fine-tuning step will not affect the performance of the final spatial model. To transfer cell labels, we followed the SCVI reference mapping workflow described in (66). Briefly, for each query spot we first identified

nearest neighbors in the reference, then assigned labels to the spot based on annotations of the neighbors. The prediction uncertainty score was calculated based on the pairwise distance between the query and its reference neighbors in the latent space. Spots with uncertainty >0.2 were labeled as 'ambiguous'.

Evaluate dimensionality reduction performance using the DLPFC dataset

All models employed expression of the top 2000 highly variable genes in each of the 12 DLPFC samples(25) as input for dimensionality reduction. For PCA, the expression was further scaled after log-normalization. Based on the ground truth layer annotation, we evaluated the quality of the obtained latent representation using two metrics: embedding consistency measured by the Silhouette score using Scikit-learn(62), and clustering accuracy measured by adjusted rank index where we clustered spots into the observed actual number of regions (cortical layers) using the R package 'Mclust'(67). The STAGATE model was configured with two graph attention layers of 128 and 30 units (i.e., 30 latent dimensions) and trained using default parameters. The SpaceFlow model had two fixed graph convolutional layers. We set the hidden dimension size 'z_dim' to 30 and again trained the model using the default setting. The baseline vanilla deep autoencoder contained two fully connected layers of dimension 128 and 30, with batch normalization and 'ELU' as the activation function. For Smoother-guided models, we constructed the regular spatial loss using the ICAR model and $\rho = 0.99$. To construct the contrastive loss, we generated 20 corrupted graphs and set the relative importance of negative samples to 0.05. When training the deep autoencoder, we also introduced an additional orthogonal loss regularizing the latent space to prevent embedding collapse.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The python package Smoother developed in this study, tutorials, and the scripts used to simulate ST data are publicly available at https://github.com/JiayuSuPKU/Smoother/ under the BSD 3-Clause License. All additional analysis codes for reproducing results and figures presented in this study will be made available as a separate GitHub repository upon the publication of this work. Descriptions and links to all public datasets analyzed in this study can be found in the corresponding Methods sections. These include the single-nucleus RNA-seq data and the 10x Visium data of mouse brain(26) downloaded from ArrayExpress (https://www.ebi.ac.uk/arrayexpress/experiments/) under accession IDs E-MTAB-11114 and E-MTAB-11115, the sci-Space mouse embryonic data(63) downloaded from GEO at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE166692, the 10x Visium invasive ductal carcinoma (IDC) data(8) with DAPI and anti-CD3 staining from 10x Genomics at https://support.10xgenomics.com/spatial-gene-expression/datasets, and the spatial-CUT&Tag mouse embryonic data(4) downloaded from GEO at https://www.ncbi.xyz/geo/query/acc.cgi?acc=GSE165217 with preprocessing scripts from Zenodo at https://zenodo.org/record/5797109#.Y9wezOyZP0p. The CRC Stereo-seq and paired

scRNA-seq data(18) were obtained from CNGB Nucleotide Sequence Archive under accession ID CNP0002432 at https://db.cngb.org/search/project/CNP0002432/. The human prostate Slide-seqV2 data (48) were obtained from https://github.com/shenglinmei/ProstateCancerAnalysis.

Competing interests

R. Rabadan is the founder of Genotwin and member of the Scientific Advisory Board of Diatech Pharmacogenetics. None of these activities are related to the results in the current manuscript.

Funding

This work was funded by the National Institutes of Health, National Cancer Institute (R35CA253126, U01CA261822, and U01CA243073 to R.R. and X.F., R37CA258829 and R21CA263381 to B.I.), the Center for Integrated Cellular Analysis (RM1HG011014 to J.S. and R.R.), and the Edward P. Evans Center for MDS at Columbia University (to J.S. and R.R.).

Authors' contributions

J.S. conceived the original idea and developed the computational framework. R.R. and D.A.K. supervised the project. J.S., J.R., X.F., G.Z., and R.S.E carried out the analysis. J.J developed the data simulation scripts. Y.W. and B.I. contributed to the analysis of MBM data and the interpretation of the results. L.A. contributed to the analysis and interpretation on human prostate cell-type prediction. J.S. took the lead in writing and revising the manuscript with input and critical feedback from all authors. All authors read and approved the final manuscript.

Acknowledgements

The authors thank Y. Choi for contributions in benchmarking, P. Sims, Y. Shen, and C. Zhang for early discussions, and members of the Rabadan lab and Knowles lab for helpful comments and feedback throughout the project.

Figure legends

Figure 1: Overview of the Smoother framework.

Smoother is a versatile and modular framework designed to incorporate spatial dependencies into various omics data analysis applications. The process initiates with the construction of a weighted spatial graph, derived from physical positions, histology, and additional features, which serves to represent spatial dependencies a priori (top). The prior is subsequently employed as a sparse loss function to regularize spatial variables, such as gene activities, cell-type compositions, and latent embeddings (bottom). Owing to its modular design, the spatial loss can be appended to pre-existing models that were initially developed for non-spatial data, potentially bridging the gap between single-cell and spatial data analysis. The Smoother toolbox includes a selection of spatially aware versions of non-spatial models, including NNLS, DWLS and SVR for cell-type deconvolution, and PCA, SCVI, and SCANVI for dimensionality reduction.

Figure 2: Construction of boundary-preserving spatial priors.

(a) Schematic illustration of the process for constructing spatial weights and covariance matrices. Smoother first builds a distance-based spatial neighborhood graph, which is subsequently pruned using histology or other features to encode boundary information and inhibit undesired cross-region interactions. **(b)** The impact of hard pruning on the covariance structure. Given the regional membership annotation, hard pruning sets the weights between adjacent spots of

different regions to zero (teal, top) or a nominal value (leak=0.01, yellow, bottom), generating distinctive correlation distributions for boundary and interior pairs of spots. **(c)** An example section (1160920F) from the human breast cancer datasets(64). Left to right: H&E staining, manual pathological annotation, and depiction of boundary and interior spots as defined by the pathological regions. **(d)** Unsupervised clustering results of the same slide in **(c).** Spots were clustered based on gene expression (left), or image-based features extracted from the H&E staining image (right). **(e)** The distribution of pairwise feature similarities between boundary pairs from two regions (orange) and interior pairs from the same region (blue), as defined in **(c)** by the pathological annotation. **(f)** The distribution of pairwise correlations imposed by the resulting spatial priors after graph pruning. Left: Soft scaling using expression- (top) and histology-based similarities (bottom). Right: Hard pruning using cluster memberships from expression (top) and histology (bottom) as defined in **(d)**.

Figure 3: Evaluation of spatial regularization effects on deconvolution accuracy using simulated data.

In the first simulation, doughnut-shaped spatial transcriptomic datasets were generated from a scRNA-seq reference(26) on a 50x50 grid. We assigned 15 cell types of high (n=5) and low (n=10) densities to overlapping spatial compartments **(a-d)**. In the second simulation, spatial transcriptomic datasets of the mouse embryo were generated from the sci-Space dataset(63) by pooling single cells of eight types barcoded at the same spot (approximately 200µm pitch) **(e-h)**. **(a)** Relative abundance of two typical cell types in a single replicate. **(b)** Deconvolution accuracy of different methods as measured by Pearson correlation, with results aggregated over ten independent replicates. Error bars denote standard errors of the mean (50 high-density and 100

low-density cell types in total). **(c)** Binary presence status (proportion >= 0.05) of two typical cell types in a single replicate. **(d)** Deconvolution accuracy of different methods, similar to **(b)**, but measured by binary prediction average precision score. **(e)** Relative abundance of the lateral plate mesoderm in one sci-Space slide. **(f)** Deconvolution accuracy of different methods as measured by Pearson correlation, with results aggregated over 14 biological slides. Error bars denote standard errors of the mean (112 cell types in total). **(g)** Binary presence status (proportion >= 0.05) of the myocyte in one sci-Space slide. **(d)** Deconvolution accuracy of different methods, similar to **(f)**, but measured by binary prediction average precision score. Smoother-guided deconvolution models: NNLS (nonnegative least square), DWLS (dampened weighted least squares, modified implementation), SVR (support vector regression), and LNR (log normal regression. The CARD model features its own implementation of spatial regularization.

Figure 4: Smoother enhances cell-type deconvolution performance in various spatial omics data.

(a-c) Detection of tumor T-cell infiltration in a 10x Visium ductal carcinoma sample(8). **(a)** Left to right: Images showing single channel CD3 immunofluorescence staining (FITC/green), T-cell proportions as estimated by Smoother-guided NNLS and CARD, respectively. **(b)** Spearman correlation of predicted T-cell proportions with CD3 fluorescence levels across all spots, with whiskers indicating the 95% confidence intervals. **(c)** Spearman correlation of predicted T-cell proportions with the expression of T cell marker genes across all spots, similar to **(b)**.

(d-g) Spatial mapping of neurons in the 10x Visium mouse brain section ST8059048(26). **(d)** Left: H&E staining. Right: UMAP visualization of the paired single-nucleus transcriptomic

reference of 52 neural subtypes. **(e)** Brain subregions identified by clustering based on estimated cell-type compositions. **(f-g)** Estimated relative abundances of cortical **(f)** and hippocampal **(g)** excitatory neurons. Color intensities are proportional to the estimated cell-type proportion and are scaled by the same factor across methods for visualization purposes. HPC: hippocampus; DG: dentate gyrus (CA4).

(**h-k**) Spatial mapping of embryonic cell types in the spatial-CUT&Tag H3K4me3 50µm data of mouse embryo(4). (**h**) Left: Overview of the spatial-CUT&Tag technology, adapted from(4). Right: T-SNE visualization of the scRNA-seq reference of 37 embryonic cell types(35). (**i**) Embryonic subregions identified by clustering based on estimated cell-type compositions. (**j**) Estimated relative abundances of five major lineages aggregated from individual cell types. Color intensities are proportional to the estimated cell-type proportion and are scaled by the same factor across methods for visualization purposes. (**k**) Standardized gene activity of selected marker genes.

Figure 5: Smoother detects tumor-specific plasma cell subtypes in colorectal adenocarcinoma Stereo-seq slide.

(a) UMAP representation of the cell types from the patient-derived paired colorectal adenocarcinoma scRNA-seq reference (18). **(b)** Expression dotplot of the top marker genes for each cell type in the scRNA-seq samples. **(c)** Pathology annotation of the histology slide, adapted from(18). **(d)** Slide subregions identified by clustering based on estimated cell-type compositions. Arrows emphasize the mucosa delineation apparent with Smoother but not CARD. **(e)** Spatial visualizations of the cell-type proportions of the three plasma cell subtypes. **(f)** B-cell receptor gene expression heatmap for the IgA plasma cell spots in the mucosa and the IgG plasma cell

voxels in the lesion, normalized across spots. **(g)** Colocalization of each cell type with the IgA and IgG plasma cells. **(h)** Gene Ontology pathway enrichment analysis of the IgA plasma cell and IgG plasma cell spots.

Figure 6: Smoother enables spatially aware joint embeddings of single-cell and Slide-seqV2 data of human prostate and improves reference mapping accuracy.

(a) Schematic overview of the spatial conversion of non-spatial auto-encoder models in the Smoother framework. Smoother enforces spatial consistency via the detachable loss function. This allows the same model to be trained and applied to both spatial and non-spatial data, generating a joint spatially aware embedding. **(b)** UMAP visualization of the latent representation of scRNA-seq data of human prostate from the Tabula Sapiens(49), colored by tissue compartment (left) and technical batch (right). The representation was generated from a pretrained RNA-only SCVI prostate model(47). **(c)** Violin plots showing the number of expressed gene (left) and total RNA counts (right) per cell or spot in data of different technologies. **(d)** UMAP visualization of the joint latent representation of the Tabula Sapiens prostate scRNA-seq reference and the Slide-seqV2 data of a healthy prostate section(48). Following the SCVI data integration workflow, the RNA-only model was fine-tuned on the query spatial data with unfrozen parameters to mitigate batch effect. **(e)** Spatial visualizations of the tissue compartment (left) and cell type prediction (right) results based on the joint RNA-only embeddings shown in **(d)**. **(f)** UMAP visualization of the joint latent representation generated by SpatialVAE. The spatially aware model has the same architecture as RNA-only models in **(b)** and **(d)**, except it was fine-tuned to minimize the proposed spatial loss in addition to the original

reconstruction and KL losses. **(g)** Spatial visualizations of the tissue compartment (left) and cell-type prediction (right) results based on the joint RNA-only embeddings shown in **(f)**.

References

1. Moffitt JR, Lundberg E, Heyn H. The emerging landscape of spatial profiling technologies. Nature Reviews Genetics. 2022.
2. Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016;353(6294):78-82.
3. Stickels RR, Murray E, Kumar P, Li JL, Marshall JL, Di Bella DJ, et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat Biotechnol. 2021;39(3):313-9.
4. Deng YX, Bartosovic M, Kukanja P, Zhang D, Liu Y, Su G, et al. Spatial-CUT&Tag: Spatially resolved chromatin modification profiling at the cellular level. Science. 2022;375(6581):681-+.
5. Deng YX, Bartosovic M, Ma S, Zhang D, Kukanja P, Xiao Y, et al. Spatial profiling of chromatin accessibility in mouse and human tissues. Nature. 2022.
6. Rao A, Barkley D, Franca GS, Yanai I. Exploring tissue architecture using spatial transcriptomics. Nature. 2021;596(7871):211-20.
7. Dries R, Zhu Q, Dong R, Eng CHL, Li HP, Liu K, et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. Genome Biol. 2021;22(1).
8. Zhao E, Stone MR, Ren X, Guenthoer J, Smythe KS, Pulliam T, et al. Spatial transcriptomics at subspot resolution with BayesSpace. Nat Biotechnol. 2021;39(11):1375-+.
9. Ma Y, Zhou X. Spatially informed cell-type deconvolution for spatial transcriptomics. Nat Biotechnol. 2022;40(9):1349-+.
10. Zhang H, Hunter MV, Chou J, Quinn JF, Zhou M, White RM, et al. BayesTME: An end-to-end method for multiscale spatial transcriptional profiling of the tissue microenvironment. Cell Systems. 2023;14(7):605-19. e7.
11. Hu J, Li XJ, Coleman K, Schroeder A, Ma N, Irwin DJ, et al. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. Nature Methods. 2021;18(11):1342-+.
12. Dong KN, Zhang SH. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. Nat Commun. 2022;13(1).
13. Ren HL, Walker BL, Cang ZX, Nie Q. Identifying multicellular spatiotemporal organization of cells with SpaceFlow. Nat Commun. 2022;13(1).
14. Wang YG, Song B, Wang SD, Chen MY, Xie Y, Xiao GH, et al. Sprod for de-noising spatially resolved transcriptomics data based on position and image information. Nature Methods. 2022;19(8):950-+.
15. Bergenstrahle L, He B, Bergenstrahle J, Abalo X, Mirzazadeh R, Thrane K, et al. Super-resolved spatial transcriptomics by deep data fusion. Nat Biotechnol. 2022;40(4):476-+.

903 16. Wei XY, Fu SL, Li HB, Liu Y, Wang S, Feng WM, et al. Single-cell Stereo-seq reveals
904 induced progenitor cells involved in axolotl brain regeneration. Science. 2022;377(6610):1062-+.
905 17. Biermann J, Melms JC, Amin AD, Wang YP, Caprio LA, Karz A, et al. Dissecting the
906 treatment-naive ecosystem of human melanoma brain metastasis. Cell. 2022;185(14):2591-+.
907 18. Zhang R, Feng Y, Ma W, Guo Y, Luo M, Li Y, et al. Spatial transcriptome unveils a
908 discontinuous inflammatory pattern in proficient mismatch repair colorectal adenocarcinoma.
909 Fundamental Research. 2022.
910 19. Rue Hv, Held L, ProQuest. Gaussian Markov random fields : theory and applications.
911 Boca Raton: Chapman & Hall/CRC; 2005.
912 20. Sardy S, Tseng P. On the statistical analysis of smoothing by maximizing dirty Markov
913 random field posterior distributions. J Am Stat Assoc. 2004;99(465):191-204.
914 21. Fan JQ, Li RZ. Variable selection via nonconcave penalized likelihood and its oracle
915 properties. J Am Stat Assoc. 2001;96(456):1348-60.
916 22. Besag J, York J, Mollie A. Bayesian Image-Restoration, with 2 Applications in Spatial
917 Statistics. Ann I Stat Math. 1991;43(1):1-20.
918 23. Wang LH, Maletic-Savatic M, Liu ZD. Region-specific denoising identifies spatial co-
919 expression patterns and intra-tissue heterogeneity in spatially resolved transcriptomics data.
920 Nat Commun. 2022;13(1).
921 24. Andreatta M, Carmona SJ. UCell: Robust and scalable single-cell gene signature scoring.
922 Comput Struct Biotec. 2021;19:3796-8.
923 25. Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, et al.
924 Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nat
925 Neurosci. 2021;24(3):425-36.
926 26. Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell2location
927 maps fine-grained cell types in spatial transcriptomics. Nat Biotechnol. 2022;40(5):661-+.
928 27. Elosua-Bayes M, Nieto P, Mereu E, Gut I, Heyn H. SPOTlight: seeded NMF regression to
929 deconvolute spatial transcriptomics spots with single-cell transcriptomes. Nucleic Acids Res.
930 2021;49(9).
931 28. Cable DM, Murray E, Zou LLS, Goeva A, Macosko EZ, Chen F, et al. Robust decomposition
932 of cell type mixtures in spatial transcriptomics. Nat Biotechnol. 2022;40(4):517-+.
933 29. Andersson A, Bergenstrahle J, Asp M, Bergenstrahle L, Jurek A, Fernandez Navarro J, et
934 al. Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography.
935 Commun Biol. 2020;3(1):565.
936 30. Dong R, Yuan GC. SpatialDWLS: accurate deconvolution of spatial transcriptomic data.
937 Genome Biol. 2021;22(1):145.
938 31. Lopez R, Li BG, Keren-Shaul H, Boyeau P, Kedmi M, Pilzer D, et al. DestVI identifies
939 continuums of cell types in spatial transcriptomics data. Nat Biotechnol. 2022.
940 32. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of
941 cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453-7.
942 33. Tsoucas D, Dong R, Chen H, Zhu Q, Guo G, Yuan GC. Accurate estimation of cell-type
943 composition from gene expression data. Nat Commun. 2019;10(1):2975.
944 34. Danaher P, Kim Y, Nelson B, Griswold M, Yang Z, Piazza E, et al. Advances in mixed cell
945 deconvolution enable quantification of cell types in spatial transcriptomic data. Nat Commun.
946 2022;13(1).

947 35. Cao JY, Spielmann M, Qiu XJ, Huang XF, Ibrahim DM, Hill AJ, et al. The single-cell
948 transcriptional landscape of mammalian organogenesis. Nature. 2019;566(7745):496-+.
949 36. Zou H, Hastie T. Regularization and variable selection via the elastic net (vol B 67, pg 301,
950 2005). J R Stat Soc B. 2005;67:768-.
951 37. Pelka K, Hofree M, Chen JH, Sarkizova S, Pirl JD, Jorgji V, et al. Spatially organized
952 multicellular immune hubs in human colorectal cancer. Cell. 2021;184(18):4734-+.
953 38. Harris RJ, Cheung A, Ng JCF, Laddach R, Chenoweth AM, Crescioli S, et al. Tumor-
954 Infiltrating B Lymphocyte Profiling Identifies IgG-Biased, Clonally Expanded Prognostic
955 Phenotypes in Triple-Negative Breast Cancer. Cancer Res. 2021;81(16):4290-304.
956 39. Gommerman JL, Rojas OL, Fritz JH. Re-thinking the functions of IgA(+) plasma cells. Gut
957 Microbes. 2014;5(5):652-62.
958 40. Xia J, Xie ZJ, Niu GM, Lu Z, Wang ZQ, Xing Y, et al. Single-cell landscape and clinical
959 outcomes of infiltrating B cells in colorectal cancer. Immunology. 2023;168(1):135-51.
960 41. Fitzerald S, O'Reilly JA, Wilson E, Joyce A, Farrell R, Kenny D, et al. Measurement of the
961 IgM and IgG Autoantibody Immune Responses in Human Serum has High Predictive Value for
962 the Presence of Colorectal Cancer. Clin Colorectal Canc. 2019;18(1):E53-E60.
963 42. Liu RX, Wen CY, Ye WB, Li YW, Chen JX, Zhang Q, et al. Altered B cell immunoglobulin
964 signature exhibits potential diagnostic values in human colorectal cancer. Iscience. 2023;26(3).
965 43. Xu YQ, Wei Z, Feng M, Zhu DX, Mei SL, Wu ZE, et al. Tumor-infiltrated activated B cells
966 suppress liver metastasis of colorectal cancers. Cell Rep. 2022;40(9).
967 44. Jasso GJ, Jaiswal A, Varma M, Laszewski T, Grauel A, Omar A, et al. Colon stroma
968 mediates an inflammation-driven fibroblastic response controlling matrix remodeling and
969 healing. Plos Biol. 2022;20(1).
970 45. Plaut E. From principal subspaces to principal components with linear autoencoders.
971 arXiv preprint arXiv:180410253. 2018.
972 46. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell
973 transcriptomics. Nature Methods. 2018;15(12):1053-+.
974 47. Gayoso A, Lopez R, Xing G, Boyeau P, Amiri VVP, Hong J, et al. A Python library for
975 probabilistic analysis of single-cell omics data. Nat Biotechnol. 2022;40(2):163-6.
976 48. Hirz T, Mei S, Sarkar H, Kfoury Y, Wu S, Verhoeven BM, et al. Dissecting the immune
977 suppressive human prostate tumor microenvironment via integrated single-cell and spatial
978 transcriptomic analyses. Nat Commun. 2023;14(1):663.
979 49. Jones RC, Karkanias J, Krasnow MA, Pisco AO, Quake SR, Salzman J, et al. The Tabula
980 Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. Science.
981 2022;376(6594):711-+.
982 50. Luecken MD, Buttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al.
983 Benchmarking atlas-level data integration in single-cell genomics. Nature Methods.
984 2022;19(1):41-+.
985 51. Yang L, Shami A. On hyperparameter optimization of machine learning algorithms:
986 Theory and practice. Neurocomputing. 2020;415:295-316.
987 52. Anselin L. Spatial econometrics : methods and models. 1st 1988. ed. Dordrecht,
988 Netherlands: Kluwer Academic Publishers; 1988.
989 53. van Erp S, Oberski DL, Mulder J. Shrinkage priors for Bayesian penalized regression. J
990 Math Psychol. 2019;89:31-50.

54. Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, et al. Squidpy: a scalable framework for spatial omics analysis. Nature Methods. 2022;19(2):171-+.
55. Lee DD, Seung HS. Algorithms for non-negative matrix factorization. Adv Neur In. 2001;13:556-62.
56. Comon P. Independent Component Analysis, a New Concept. Signal Process. 1994;36(3):287-314.
57. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. Advances in Neural Information Processing Systems 32 (Nips 2019). 2019;32.
58. Diamond S, Boyd S. CVXPY: A Python-Embedded Modeling Language for Convex Optimization. J Mach Learn Res. 2016;17.
59. Yuan ZY, Pan WT, Zhao X, Zhao FY, Xu ZM, Li X, et al. SODB facilitates comprehensive exploration of spatial omics data. Nature Methods. 2023;20(3):387-+.
60. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19.
61. Hao YH, Hao S, Andersen-Nissen E, Mauck WM, Zheng SW, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573-+.
62. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825-30.
63. Srivatsan SR, Regier MC, Barkan E, Franks JM, Packer JS, Grosjean P, et al. Embryo-scale, single-cell spatial transcriptomics. Science. 2021;373(6550):111-+.
64. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. Nat Genet. 2021;53(9):1334-47.
65. Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HWY, et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis (vol 53, pg 403, 2021). Nat Genet. 2021;53(6):935-.
66. Sikkema L, Strobl DC, Zappia L, Madissoon E, Markov NS, Zaragosi L-E, et al. An integrated cell atlas of the human lung in health and disease. bioRxiv. 2022:2022.03. 10.483747.
67. Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. The R journal. 2016;8(1):289.

Figure 1

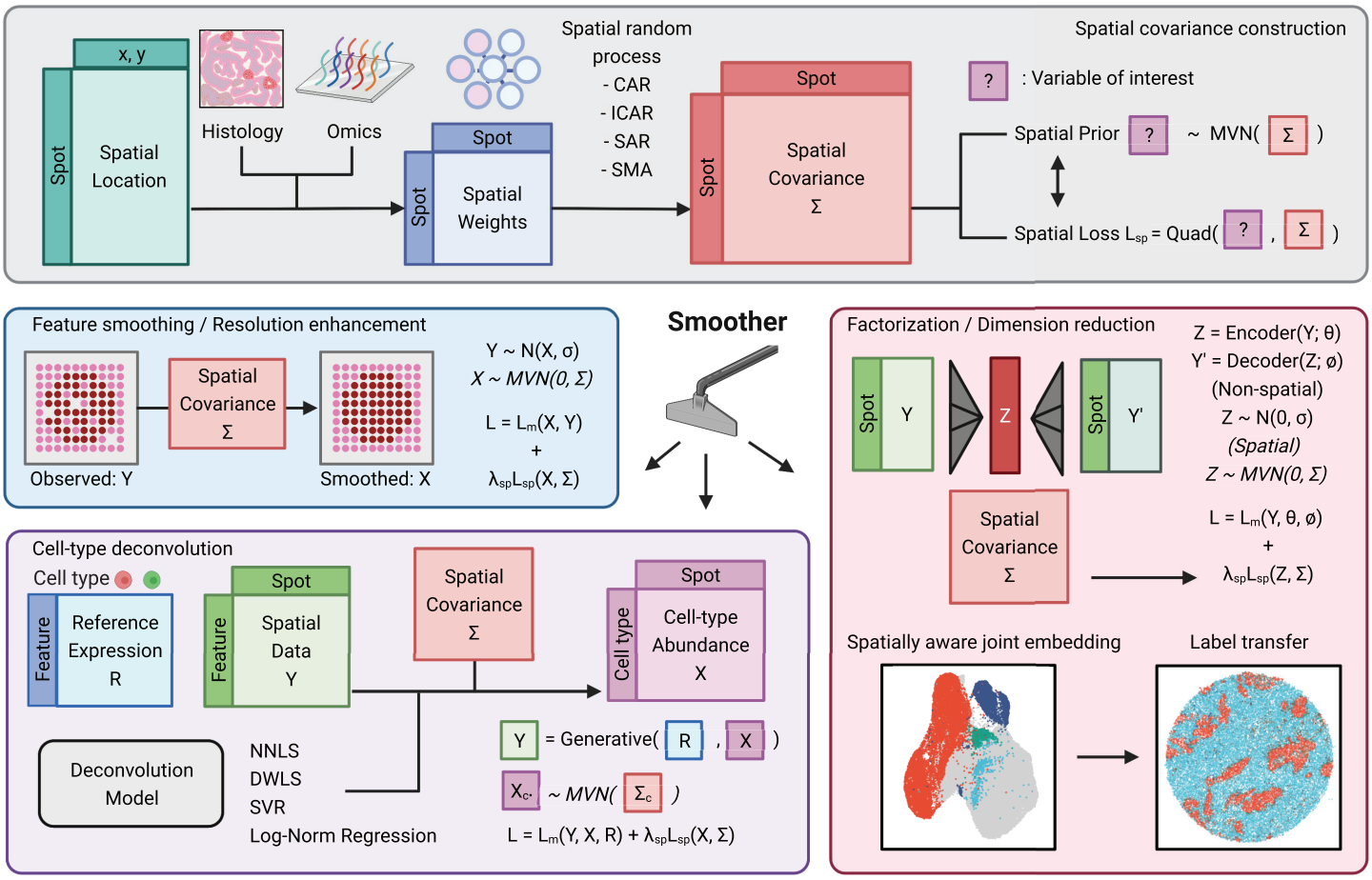

Spatial covariance construction
x, y
Spot
Spatial Location
Histology
Omics
Spot
Spot
Spatial Weights
Spatial random process
- CAR
- ICAR
- SAR
- SMA
Spot
Spot
Spatial Covariance Σ
? : Variable of interest
Spatial Prior ? ~ MVN(Σ)
Spatial Loss Lsp = Quad(? , Σ)
Feature smoothing / Resolution enhancement
Spatial Covariance Σ
Observed: Y
Smoothed: X
Y ~ N(X, σ)
X ~ MVN(0, Σ)
L = Lm(X, Y) + λspLsp(X, Σ)
Smoother
Factorization / Dimension reduction
Spot
Y
Z
Spot
Y'
Z = Encoder(Y; θ)
Y' = Decoder(Z; ø)
(Non-spatial)
Z ~ N(0, σ)
(Spatial)
Z ~ MVN(0, Σ)
Spatial Covariance Σ
L = Lm(Y, θ, ø) + λspLsp(Z, Σ)
Spatially aware joint embedding
Label transfer
Cell-type deconvolution
Cell type
Feature
Reference Expression R
Spot
Feature
Spatial Data Y
Spatial Covariance Σ
Spot
Cell type
Cell-type Abundance X
Deconvolution Model
NNLS
DWLS
SVR
Log-Norm Regression
Y = Generative(R , X)
Xc· ~ MVN(Σc)
L = Lm(Y, X, R) + λspLsp(X, Σ)

Figure 2

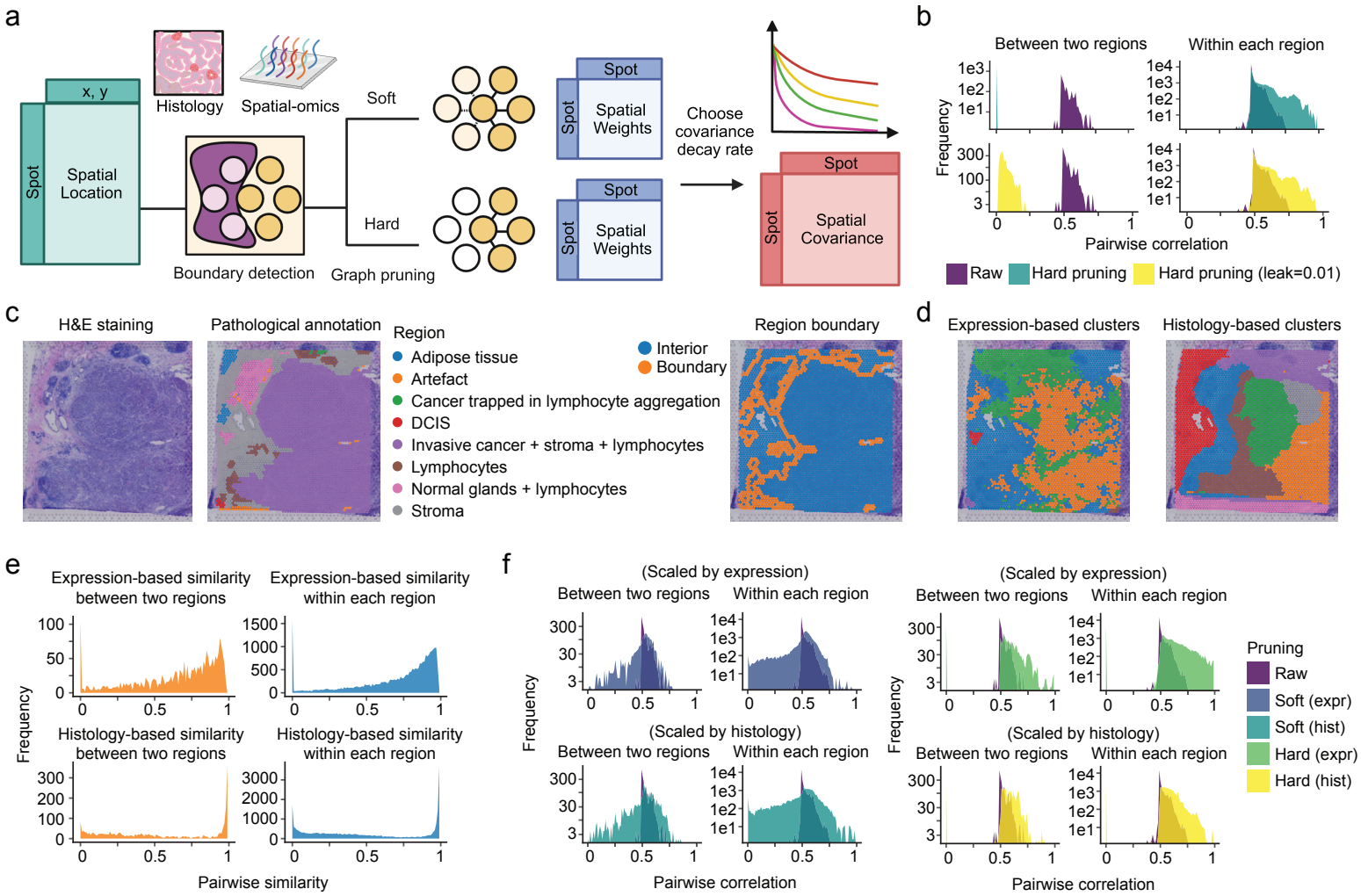

a
x, y
Spot
Spatial Location
Histology
Spatial-omics
Boundary detection
Soft
Hard
Graph pruning
Spot
Spot
Spatial Weights
Spot
Spot
Spatial Weights
Choose covariance decay rate
Spot
Spot
Spatial Covariance
b
Between two regions
Within each region
Frequency
Pairwise correlation
Raw
Hard pruning
Hard pruning (leak=0.01)
c
H&E staining
Pathological annotation
Region
Adipose tissue
Artefact
Cancer trapped in lymphocyte aggregation
DCIS
Invasive cancer + stroma + lymphocytes
Lymphocytes
Normal glands + lymphocytes
Stroma
Region boundary
Interior
Boundary
d
Expression-based clusters
Histology-based clusters
e
Expression-based similarity between two regions
Expression-based similarity within each region
Histology-based similarity between two regions
Histology-based similarity within each region
Frequency
Pairwise similarity
f
(Scaled by expression)
Between two regions
Within each region
(Scaled by histology)
Frequency
Pairwise correlation
Pruning
Raw
Soft (expr)
Soft (hist)
Hard (expr)
Hard (hist)

Figure 3

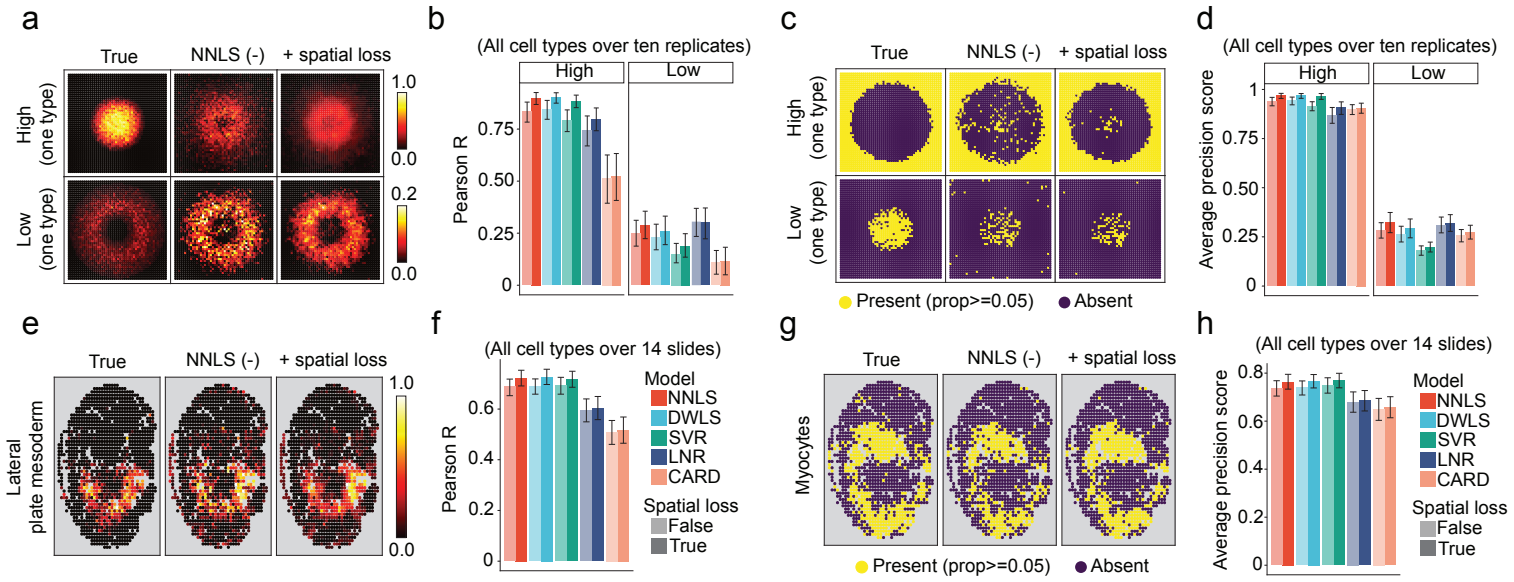

a
True
NNLS (-)
+ spatial loss
High (one type)
Low (one type)
1.0
0.0
0.2
0.0
b
(All cell types over ten replicates)
High
Low
Pearson R
0.75
0.50
0.25
0
c
True
NNLS (-)
+ spatial loss
High (one type)
Low (one type)
Present (prop>=0.05)
Absent
d
(All cell types over ten replicates)
High
Low
Average precision score
1
0.75
0.50
0.25
0
e
True
NNLS (-)
+ spatial loss
Lateral plate mesoderm
1.0
0.0
f
(All cell types over 14 slides)
Pearson R
0.6
0.4
0.2
0
Model
NNLS
DWLS
SVR
LNR
CARD
Spatial loss
False
True
g
True
NNLS (-)
+ spatial loss
Myocytes
Present (prop>=0.05)
Absent
h
(All cell types over 14 slides)
Average precision score
0.8
0.6
0.4
0.2
0
Model
NNLS
DWLS
SVR
LNR
CARD
Spatial loss
False
True

Figure 4

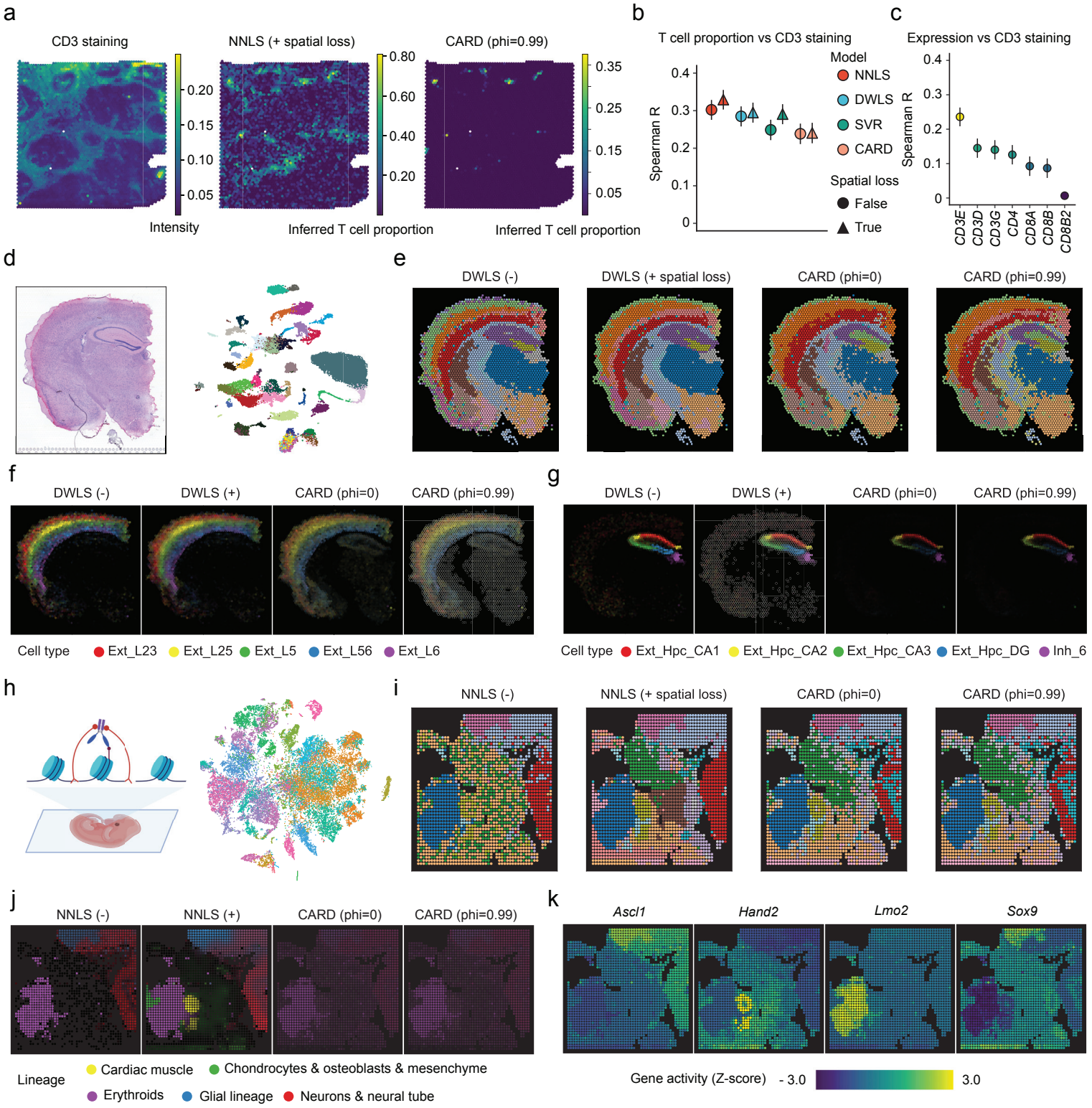

a
CD3 staining
NNLS (+ spatial loss)
CARD (phi=0.99)
0.20
0.15
0.10
0.05
Intensity
0.80
0.60
0.40
0.20
Inferred T cell proportion
0.35
0.25
0.15
0.05
Inferred T cell proportion
b
T cell proportion vs CD3 staining
Spearman R
0.4
0.3
0.2
0.1
0
Model
NNLS
DWLS
SVR
CARD
Spatial loss
False
True
c
Expression vs CD3 staining
Spearman R
0.4
0.3
0.2
0.1
0
CD3E
CD3D
CD3G
CD4
CD8A
CD8B
CD8B2
d
e
DWLS (-)
DWLS (+ spatial loss)
CARD (phi=0)
CARD (phi=0.99)
f
DWLS (-)
DWLS (+)
CARD (phi=0)
CARD (phi=0.99)
Cell type
Ext_L23
Ext_L25
Ext_L5
Ext_L56
Ext_L6
g
DWLS (-)
DWLS (+)
CARD (phi=0)
CARD (phi=0.99)
Cell type
Ext_Hpc_CA1
Ext_Hpc_CA2
Ext_Hpc_CA3
Ext_Hpc_DG
Inh_6
h
i
NNLS (-)
NNLS (+ spatial loss)
CARD (phi=0)
CARD (phi=0.99)
j
NNLS (-)
NNLS (+)
CARD (phi=0)
CARD (phi=0.99)
Lineage
Cardiac muscle
Chondrocytes & osteoblasts & mesenchyme
Erythroids
Glial lineage
Neurons & neural tube
k
Ascl1
Hand2
Lmo2
Sox9
Gene activity (Z-score)
- 3.0
3.0

Figure 5

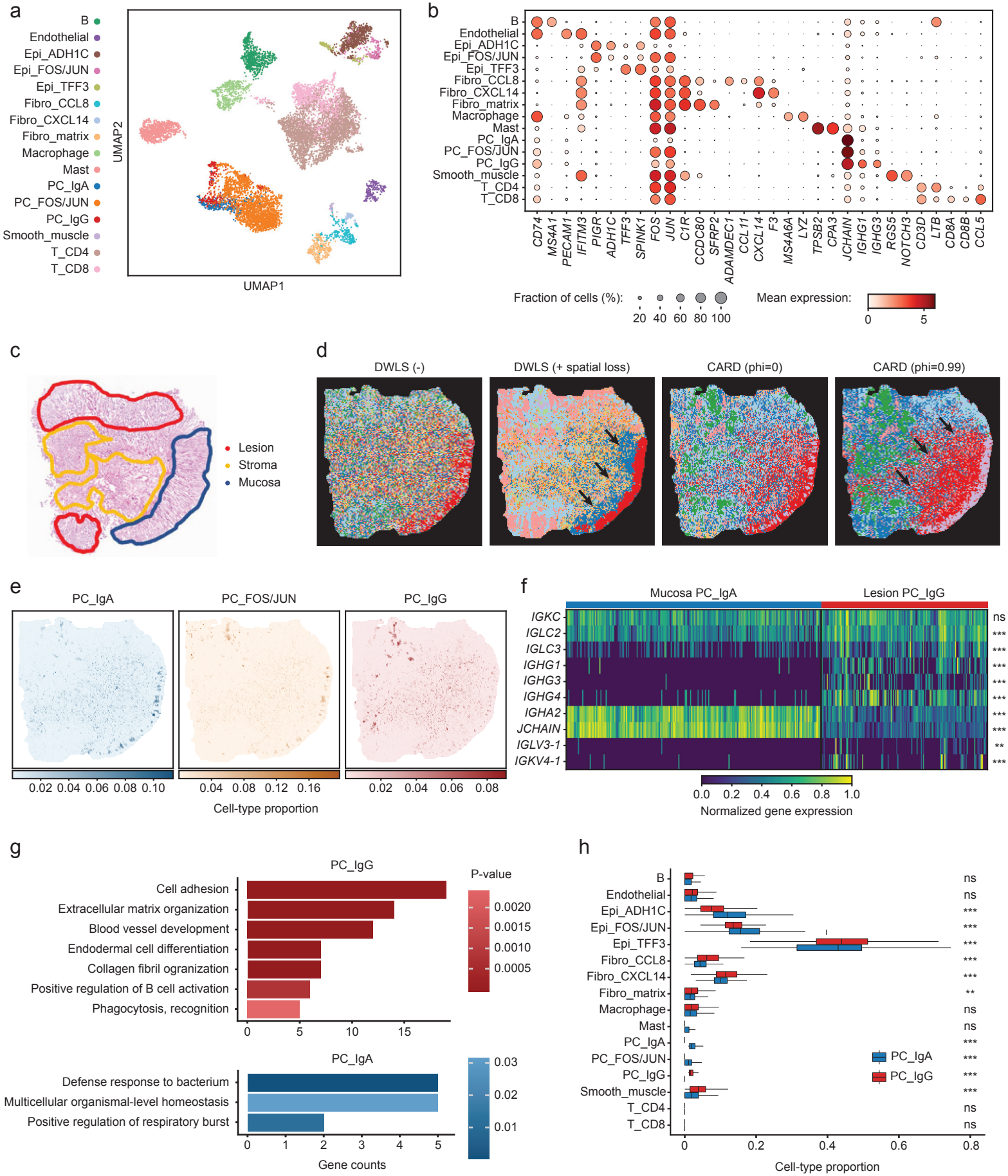
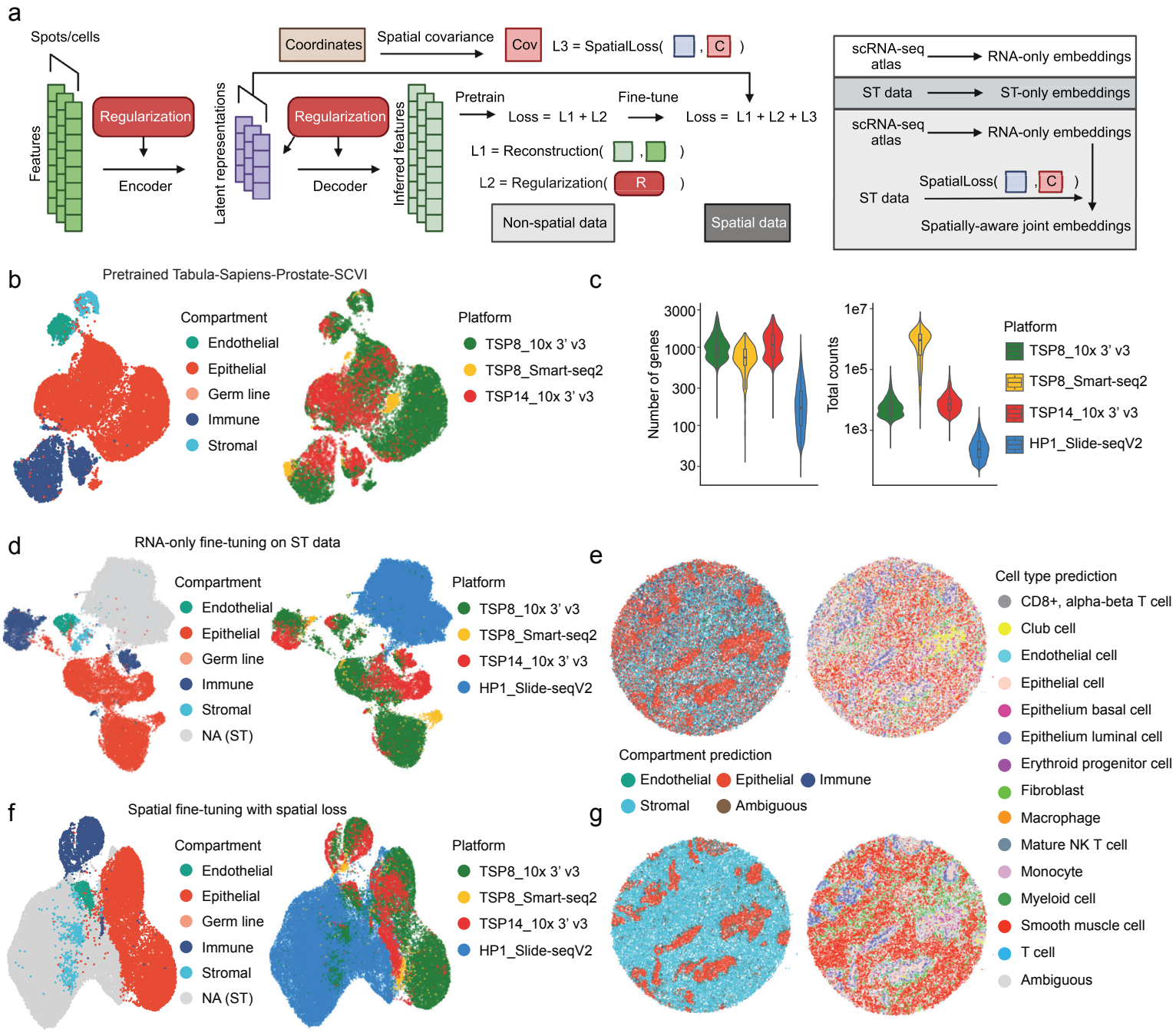

a
B
Endothelial
Epi_ADH1C
Epi_FOS/JUN
Epi_TFF3
Fibro_CCL8
Fibro_CXCL14
Fibro_matrix
Macrophage
Mast
PC_IgA
PC_FOS/JUN
PC_IgG
Smooth_muscle
T_CD4
T_CD8
UMAP2
UMAP1
b
CD74 MS4A1 PECAM1 IFITM3 PIGR ADH1C TFF3 SPINK1 FOS JUN C1R CCDC80 SFRP2 ADAMDEC1 CCL11 CXCL14 F3 MS4A6A LYZ TPSB2 CPA3 JCHAIN IGHG1 IGHG3 RGS5 NOTCH3 CD3D LTB CD8A CD8B CCL5
Fraction of cells (%):
20 40 60 80 100
Mean expression:
0 5
c
Lesion
Stroma
Mucosa
d
DWLS (-)
DWLS (+ spatial loss)
CARD (phi=0)
CARD (phi=0.99)
e
PC_IgA
PC_FOS/JUN
PC_IgG
0.02 0.04 0.06 0.08 0.10
0.04 0.08 0.12 0.16
0.02 0.04 0.06 0.08
Cell-type proportion
f
Mucosa PC_IgA
Lesion PC_IgG
IGKC ns
IGLC2 ***
IGLC3 ***
IGHG1 ***
IGHG3 ***
IGHG4 ***
IGHA2 ***
JCHAIN ***
IGLV3-1 **
IGKV4-1 ***
0.0 0.2 0.4 0.6 0.8 1.0
Normalized gene expression
g
PC_IgG
Cell adhesion
Extracellular matrix organization
Blood vessel development
Endodermal cell differentiation
Collagen fibril ogranization
Positive regulation of B cell activation
Phagocytosis, recognition
0 5 10 15
P-value
0.0020
0.0015
0.0010
0.0005
PC_IgA
Defense response to bacterium
Multicellular organismal-level homeostasis
Positive regulation of respiratory burst
0 1 2 3 4 5
Gene counts
0.03
0.02
0.01
h
B ns
Endothelial ns
Epi_ADH1C ***
Epi_FOS/JUN ***
Epi_TFF3 ***
Fibro_CCL8 ***
Fibro_CXCL14 ***
Fibro_matrix **
Macrophage ns
Mast ns
PC_IgA ***
PC_FOS/JUN ***
PC_IgG ***
Smooth_muscle ***
T_CD4 ns
T_CD8 ns
PC_IgA
PC_IgG
0 0.2 0.4 0.6 0.8
Cell-type proportion

Figure 6


a
Spots/cells
Features
Regularization
Encoder
Latent representations
Regularization
Decoder
Inferred features
Coordinates
Spatial covariance
Cov
L3 = SpatialLoss(, C)
Pretrain
Loss = L1 + L2
Fine-tune
Loss = L1 + L2 + L3
L1 = Reconstruction(,)
L2 = Regularization(R)
Non-spatial data
Spatial data
scRNA-seq atlas
RNA-only embeddings
ST data
ST-only embeddings
scRNA-seq atlas
RNA-only embeddings
ST data
SpatialLoss(, C)
Spatially-aware joint embeddings
b
Pretrained Tabula-Sapiens-Prostate-SCVI
Compartment
Endothelial
Epithelial
Germ line
Immune
Stromal
Platform
TSP8_10x 3' v3
TSP8_Smart-seq2
TSP14_10x 3' v3
c
Number of genes
3000
1000
300
100
30
Total counts
1e7
1e5
1e3
Platform
TSP8_10x 3' v3
TSP8_Smart-seq2
TSP14_10x 3' v3
HP1_Slide-seqV2
d
RNA-only fine-tuning on ST data
Compartment
Endothelial
Epithelial
Germ line
Immune
Stromal
NA (ST)
Platform
TSP8_10x 3' v3
TSP8_Smart-seq2
TSP14_10x 3' v3
HP1_Slide-seqV2
e
Compartment prediction
Endothelial
Epithelial
Immune
Stromal
Ambiguous
Cell type prediction
CD8+, alpha-beta T cell
Club cell
Endothelial cell
Epithelial cell
Epithelium basal cell
Epithelium luminal cell
Erythroid progenitor cell
Fibroblast
Macrophage
Mature NK T cell
Monocyte
Myeloid cell
Smooth muscle cell
T cell
Ambiguous
f
Spatial fine-tuning with spatial loss
Compartment
Endothelial
Epithelial
Germ line
Immune
Stromal
NA (ST)
Platform
TSP8_10x 3' v3
TSP8_Smart-seq2
TSP14_10x 3' v3
HP1_Slide-seqV2
g